Validated Chebyshev-based computations for ordinary and partial differential equations

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan de Vrije Universiteit Amsterdam, op gezag van de rector magnificus prof.dr. V. Subramaniam, in het openbaar te verdedigen ten overstaan van de promotiecommissie van de Faculteit der Bètawetenschappen op woensdag 11 april 2018 om 13.45 uur in de aula van de universiteit, De Boelelaan 1105

doorgeboren te Naarden
promotoren: prof.dr. G.J.B. van den Berg
prof.dr. B.W. Rink
# Contents

1 Introduction ...................................................... 1  
1.1 Dynamical systems ........................................ 1  
1.1.1 Qualitative dynamical system theory .................. 2  
1.1.2 Validated computations for dynamical systems .......... 2  
1.1.3 Non-rigorous numerical methods ......................... 3  
1.2 Validated numerical methods ................................ 5  
1.2.1 Interval arithmetic ...................................... 6  
1.2.2 A verification method based on Newton’s method .... 8  
1.3 Example: Chebyshev series for delay differential equations . 13  
1.4 Outline of this thesis ....................................... 19  
1.5 Conclusions and future work ............................... 22  

2 Rigorous numerics for ODEs using Chebyshev series and domain decomposition 25  
2.1 Introduction .................................................. 25  
2.2 Preliminaries ............................................... 31  
2.2.1 Chebyshev series ........................................ 31  
2.2.2 Sequence spaces ......................................... 33  
2.3 Rigorous numerics for periodic orbits ..................... 33  
2.3.1 Chebyshev operator for periodic orbits ............... 34  
2.3.2 Finite Dimensional Reduction .......................... 37  
2.3.3 A Newton-like scheme .................................. 37  
2.4 Domain decomposition .................................... 40  
2.4.1 A heuristic procedure for computing $\mathcal{P}_{mn}$ .... 41  
2.4.2 Approximation of the complex singularities .......... 42  
2.5 Applications: periodic and heteroclinic orbits in the Lorenz system 44  
2.5.1 Main algorithm .......................................... 44  
2.5.2 Periodic orbit on the Lorenz-attractor ............... 45  
2.5.3 Family of periodic orbits near a homoclinic connection . 47  
2.5.4 Heteroclinic orbit ...................................... 51  
2.6 The estimates needed to prove contraction ............... 53  
2.6.1 Computation of the $Y$-bounds ......................... 54  
2.6.2 Computation of the $Z$-bounds ......................... 54
2.6.3 Modifications for non-periodic boundary conditions

3 Validated computations for connecting orbits in polynomial vector fields

3.1 Introduction

3.2 Preliminaries

3.2.1 Chebyshev series

3.2.2 Geometrically decaying sequences

3.2.3 Multivariate sequences

3.3 An equivalent zero finding problem

3.3.1 Charts on the (un)stable manifolds

3.3.2 Chebyshev series and domain decomposition

3.3.3 The connecting orbit map

3.4 Functional analytic setup

3.4.1 Projections

3.4.2 An equivalent fixed-point problem

3.5 Bounds for proving contraction

3.5.1 Y-bounds

3.5.2 Z-bounds

3.6 Applications: traveling fronts in parabolic PDEs

3.6.1 Lotka-Volterra

3.6.2 Traveling fronts in a fourth order parabolic PDE

4 Validated integration of semilinear parabolic PDEs

4.1 Introduction

4.2 Preliminaries

4.2.1 Sequence spaces

4.2.2 Chebyshev interpolation

4.3 Functional analytic setup

4.3.1 An equivalent zero-finding problem

4.3.2 Finite dimensional reduction

4.3.3 A posteriori analysis

4.4 General estimates

4.5 Y-bounds

4.6 Z-bounds

4.6.1 First order bounds

4.6.2 Second order bounds

4.7 Implementation details

4.8 Long time integration

4.8.1 Domain decomposition

4.8.2 Time stepping

4.9 Applications

4.9.1 Fisher’s equation

4.9.2 Swift-Hohenberg
1 Introduction

1.1 Dynamical systems

Dynamical system theory is the branch of mathematics which studies the time evolution of quantities which change according to a prescribed set of rules. Many fundamental laws in physics are governed by such systems. For example, in elementary mechanics the motion of a point mass is governed by Newton’s second law, which states that its acceleration is proportional to the force acting on it. The quantities which change in time in this scenario are the position and velocity vectors of the point mass. The associated evolution rule is given by a second order ordinary differential equation (ODE). In particular, the motion of the particle is completely determined by its initial position and velocity. Furthermore, in this specific example, the quantities which evolve in time (the position and velocity vectors) are elements of a six dimensional (vector) space, which is commonly referred to as the phase space of the dynamical system.

Another example from physics is the dynamical system generated by the so-called heat equation, which is derived from Fourier’s law and models the temperature of a solid three dimensional body $\Omega \subset \mathbb{R}^3$ as a function of time and space. In this case, the quantity which changes in time is the temperature of the body, i.e., a real-valued function on $\Omega$. The associated evolution rule is given by a partial differential equation (PDE). A fundamental difference with the previous example is that in this scenario the phase space, i.e., the space in which the evolving quantity is contained, consists of (suitably chosen) real-valued functions on $\Omega$ and is thus infinite dimensional.

In general, a dynamical system consists of a triple $(X, T, \varphi)$, where $X$ is usually a topological space with some additional structure (e.g. $X$ could be a Banach space or a manifold) and is referred to as the phase space, $T \in \{\mathbb{Z}, \mathbb{N}_0, \mathbb{R}, \mathbb{R}_{\geq 0}\}$ is an ordered set which is interpreted as time, and $\varphi : T \times X \rightarrow X$ is an evolution rule which governs the “motion” of elements in $X$ in a continuous way. The evolution rule $\varphi$, which is more commonly referred to as a flow, is required to specify the trajectory or orbit $\{\varphi(t, x) : t \in T\}$ of a point $x$ uniquely and should leave points fixed at $t = 0$. Formally, this corresponds to requiring that $\varphi(t, \varphi(s, x)) = \varphi(t + s, x)$ and $\varphi(0, x) = x$ for all $t, s \in T$ and $x \in X$. While general dynamical system theory deals with rather abstract situations, in this dissertation we are exclusively concerned with continuous dynamical systems (i.e. $T \in \{\mathbb{R}, \mathbb{R}_{\geq 0}\}$) generated by
In most real-world applications the physical phenomena and corresponding models are inherently nonlinear. This hampers a detailed study of the associated dynamics; while linear differential equations admit explicit solution formulae, it is usually impossible to find closed form expressions for solutions of nonlinear equations. In applied settings, this issue arises already for seemingly “simple” problems. For example, the famous three-body problem, which models the motion of three point masses in three dimensional space under gravitational force, does not admit explicit solution formulae. Poincaré proved this at the end of the nineteenth century and raised the following question: how can one study nonlinear differential equations if no explicit solutions are available? This question led to the development of modern dynamical system theory as we know it today.

The central idea in dynamical system theory is to study qualitative dynamical properties instead of explicit solution formulae. An example of a qualitative property of a dynamical system is the characterization of the asymptotic behavior of a typical orbit; does a generic orbit tend to oscillatory behavior, or to a steady state, etc. To answer questions like these, one typically starts by studying local dynamical properties, e.g., the existence of steady states and the dynamics in a small neighborhood around them (which can be studied through linearization). A more challenging task is the study of global dynamical objects, such as periodic orbits, invariant manifolds and connecting orbits between them.

Altogether, the local dynamics and (non-)existence of global objects can be used as a framework to establish more complicated dynamical phenomena through forcing theorems. A well-known example is Shilnikov’s theorem, which states that under certain (local) conditions, the existence of a homoclinic orbit (a global object) enforces the existence of an infinite number of periodic orbits nearby. In general, forcing theorems provide valuable insight into global dynamical behavior and can answer questions such as whether a dynamical system is chaotic or not.

The study of local dynamics is often still feasible with pen and paper for relatively simple and low dimensional problems. However, the analysis of global objects, such as the ones used in typical forcing theorems, is difficult (if not impossible) to perform by hand. For this reason, one often resorts to numerical methods for studying global dynamics. While numerical methods are extremely useful tools for gaining valuable insight into complicated dynamics, which would otherwise be impossible to obtain with merely a pen and paper analysis, the numerical results produced by computers are not mathematically rigorous. In particular, numerical results cannot be used as ingredients for forcing theorems. We will explain in the next section in what sense standard numerical methods are not mathematically rigorous. For now, it suffices to say that numerical methods “simulate” the true
problem (similar to simulating a physical experiment) and do not solve the exact mathematical problem.

The objective of this dissertation is to bridge the gap between standard numerical methods and mathematical analysis by developing validated numerical methods for dynamical systems. In short, a validated numerical method is a theorem, in the ordinary mathematical sense, whose hypotheses can be verified with the aid of a computer. In particular, the results of a validated numerical method can be directly interpreted in a mathematically rigorous sense. This is to be contrasted with theorems in numerical analysis, which guarantee the convergence of a numerical algorithm provided the solution of a problem exists, as some computational parameter (e.g. a discretization size) goes to zero. While such theorems are extremely useful for establishing the reliability of a numerical method, in mathematical analysis the existence of a solution is the very statement we are trying to prove. The goal of a validated numerical method is to establish the existence of a true solution directly from a numerical simulation. In addition, a validated numerical method provides a rigorous upper bound for the error between the exact and approximate solution, which traditional numerical methods usually do not provide.

The development of validated numerical methods is an active field and is not restricted to dynamical system theory. It is outside the scope of this text to give a comprehensive overview of the field. That being said, we mention three seminal papers in the history of computer-assisted proofs for dynamical systems: in [45] the universality of the Feigenbaum constant is established, in [55] it is shown that the Lorenz system is chaotic, and finally in [69] the existence of the Lorenz attractor is proven.

1.1.3 Non-rigorous numerical methods

In this section we describe a popular and effective strategy used in the development of many numerical (and mathematical) methods. The validated numerical methods developed in this thesis are all based on this strategy. The main idea is to characterize the object of interest, e.g. a periodic or heteroclinic orbit, as a zero of a mapping \( F : \mathcal{X} \to \mathcal{Y} \), where \( \mathcal{X} \) and \( \mathcal{Y} \) are suitably chosen vector spaces. To ensure that the zero finding problem is well-posed, the map \( F \) needs to be constructed in such a way that the zeros of interest are locally unique. The basic principle of many numerical methods is then to construct a finite dimensional approximation of \( F \), which is amenable to numerical computations, and to approximate its zeros by using Newton’s method.

**Example 1.1.1** (Periodic orbits of finite dimensional ODEs). Suppose we wish to compute a periodic orbit \( u \) of a smooth nonlinear vectorfield \( g : \mathbb{R}^n \to \mathbb{R}^n \). A periodic orbit \( u : [0, L] \to \mathbb{R}^n \) of \( g \) is characterized by

\[
\begin{aligned}
\frac{du}{dt} &= g(u), & t \in [0, L], \\
u(0) &= u(L),
\end{aligned}
\]
where \( L > 0 \) is the (unknown) period. Since we seek for periodic orbits, we expand \( u \) in the Fourier basis on \([0, L]\):

\[
u(t) = \sum_{k \in \mathbb{Z}} a_k e^{i k \omega t}, \quad \omega = \frac{2\pi}{L}.
\]

Here \( \{a_k \in \mathbb{C}^n : k \in \mathbb{Z}\} \) are the unknown Fourier coefficients of \( u \), which need to satisfy \( \overline{a_k} = a_{-k} \) for all \( k \in \mathbb{Z} \), since \( u \) is real-valued.

Substitution of the Fourier expansion of \( u \) into (1.1) yields an infinite dimensional system of algebraic equations for the angular frequency \( \omega \) and the Fourier coefficients \((a_k)_{k \in \mathbb{Z}}\):

\[
i \omega k a_k - c_k(a) = 0, \quad k \in \mathbb{Z}.
\]

Here \((c_k(a))_{k \in \mathbb{Z}}\) are the Fourier coefficients of the map

\[
t \mapsto g \left( \sum_{k \in \mathbb{Z}} a_k e^{i k \omega t} \right).
\]

Therefore, solutions of (1.1) can be characterized as zeros of the map

\[
(\omega, a) \mapsto [i \omega k a_k - c_k(a)]_{k \in \mathbb{Z}},
\]

which satisfy the symmetry relation \( \overline{a_k} = a_{-k} \). There are many choices for the domain (and range) of this map. Regardless of the specific choices, (1.2) induces a mapping between infinite dimensional vector spaces.

Zeros of (1.2) are not locally unique, since solutions of (1.1) are invariant under translations in time. To get rid of this translation symmetry, we need to introduce a phase condition to fix the time parameterization of the orbit. For example, one could require that the periodic orbit goes through a prescribed hyperplane (a Poincaré section) at \( t = 0 \) by appending the equation

\[
\langle v, u(0) - u_0 \rangle = 0, \quad u_0, v \in \mathbb{R}^n,
\]

to the system in (1.2). Here \( u_0 \approx u(0) \) and \( v \approx u'(0) \) are fixed (numerically determined) vectors which define a hyperplane.

**Remark 1.1.2.** In this example the domain and range of \( F : \mathcal{X} \rightarrow \mathcal{Y} \) are infinite dimensional sequence spaces. For a standard numerical method the exact choices of \( \mathcal{X} \) and \( \mathcal{Y} \) (e.g. the choice of a norm) do no matter. For a validated numerical method, however, the choice of \( \mathcal{X} \) is extremely important, as it corresponds to the choice of a functional analytic framework in which the numerical results are to be interpreted (in a mathematically rigorous manner).

Let us henceforth assume that \( \mathcal{X} \) and \( \mathcal{Y} \) are infinite dimensional Banach spaces. In practice, to define a finite dimensional approximation of \( F \), we first construct finite dimensional projections \( \Pi_{\mathcal{X}, m} : \mathcal{X} \rightarrow \mathcal{X}_m \) and \( \Pi_{\mathcal{Y}, m} : \mathcal{Y} \rightarrow \mathcal{Y}_m \), where
$X_m \subset\mathcal{X}$ and $Y_m \subset\mathcal{Y}$ are finite dimensional subspaces of dimension $m \in \mathbb{N}$. The choices of the projections correspond to the specific approximation techniques one wants to use and depend heavily on the set up of the zero finding problem. For instance, in Example 1.1.1 we would simply define the projections by truncating the number of Fourier coefficients, i.e.,

$$\Pi_{X,m}(\omega, a) := \left[\omega a_{-K} \ldots a_K\right]^T,$$

where $K \in \mathbb{N}$ is a fixed truncation parameter. In this case $m = 2(K+1)$ and $\Pi_{Y,m}$ is defined in an analogous manner.

To emphasize the relationship between the choice of the mapping $F$, the finite dimensional projections and corresponding numerical approximation schemes, note that we could also have characterized solutions of (1.1) by integrating the equation and setting up a zero finding map on a suitably chosen function space (the continuous function on $[0, L]$ for example). In this case, the finite dimensional projections could correspond to linear spline approximations or high-order polynomial interpolations. Yet another approach would be to construct a zero finding map by using approximation techniques based on other orthogonal bases than the trigonometric polynomials (e.g. Legendre or Chebyshev polynomials).

After the finite dimensional projections have been chosen, we construct a finite dimensional approximation $F_m : X_m \to Y_m$ of $F$ by

$$F_m := \Pi_{Y,m} \circ F \circ \Pi_{X,m}|_{X_m}.$$ 

Finally, we use a computer to determine an approximate zero $\hat{x} \in X_m$ of $F_m$. The main mathematical question which arises at this point is the following: does the existence of an approximate zero $\hat{x}$ of $F_m$ imply the existence of a true zero $x^*$ of $F$? If so, how close is the approximate zero to the exact one? These questions remain usually unanswered in traditional numerical methods.

### 1.2 Validated numerical methods

The field of validated numerics, commonly also referred to as rigorous numerics, is concerned with the development of mathematical theorems whose hypotheses can be verified with the aid of a computer. A successful application of such a theorem is often referred to as a computer-assisted proof or CAP for short. A CAP involves the construction of a carefully defined mathematical framework, tailored to the problem at hand, in which errors induced by numerical approximations can be analyzed in a mathematically rigorous manner. The approximation errors one typically tries to manage are: (i) rounding errors which arise from the fact that computers can only represent a finite subset of the real numbers, (ii) truncation or discretization errors which are caused by replacing the exact problem with an approximate one.

The first issue is related to how the computer performs arithmetic (and not to the mathematical problem itself) and can be solved by using interval arithmetic,
see Section 1.2.1. The second issue is highly nontrivial for nonlinear problems and requires a careful analysis on paper. The main objective in this dissertation is to develop mathematical methods for analyzing problems of the second type. The general strategy used in this thesis to analyze such problems is outlined in Section 1.2.2.

1.2.1 Interval arithmetic

In this section we briefly address the issue of how one can infer mathematically rigorous statements from numerical computations on a computer. We start by recalling the notion of floating point numbers and then introduce the basic principles of interval arithmetic. The explanation is only meant to serve as a brief introduction and many details are omitted. The interested reader is referred to [58,65,70] for a more comprehensive introduction.

Since computers have a limited amount of memory, they can only store a finite subset of the real numbers. Therefore, on a computer most real numbers are approximated and do not have an exact representation. The standard arithmetic on the computer, however, does not keep track of the approximation errors (which are more commonly referred to as rounding errors). For this reason, the results from standard computations on a computer cannot be used in mathematical arguments. The idea of interval arithmetic is to overcome this problem by simply keeping track of the rounding errors. To explain how this works, we first give a rough explanation of how computers approximate real numbers and perform arithmetic on them.

Let $\beta \in \mathbb{N}_{\geq 2}$ and recall that any number $x \in \mathbb{R}$ can be written in the form

$$x = \text{sign}(x) \sum_{j=0}^{\infty} x_j \beta^{q-j}, \quad x_j \in \{0, \ldots, \beta - 1\},$$

where $q \in \mathbb{Z}$, which is often abbreviated as

$$x = \text{sign}(x) \cdot \beta^q \cdot (x_0.x_1x_2\ldots)_\beta. \quad (1.3)$$

The representation in (1.3) is unique if one imposes the constraints that $x_0 \neq 0$, i.e., the exponent $q$ is minimal, and that for every $k \in \{l \in \mathbb{N}_0 : x_l = \beta - 1\}$ there exists a $j > k$ such that $x_j < \beta - 1$. In this case, the sequence $(x_0.x_1x_2\ldots)_\beta$ is referred to as the mantissa of $x$ with respect to the basis $\beta$. Modern computers use the representation in (1.3) to store a finite subset of the real numbers of the form

$$\mathbb{F}_{q_{\min},q_{\max}}^{p,\beta} := \left\{(-1)^\sigma \cdot \beta^q \cdot (x_0.x_1x_2\ldots x_{p-1})_\beta : \sigma \in \{\pm 1\}, \quad 0 \leq x_j \leq \beta_j - 1 \right\},$$

where $q_{\min}, q_{\max} \in \mathbb{Z}$ and $p \in \mathbb{N}$. The elements in $\mathbb{F}_{q_{\min},q_{\max}}^{p,\beta}$ are referred to as floating point numbers. The number $p$ is called the precision of the floating point system $\mathbb{F}_{q_{\min},q_{\max}}^{p,\beta}$. In practice, computers also store special symbols for $\pm \infty$, which
provide a systematic way for dealing with numbers that are too “large” (i.e. they are bigger than the largest floating point number in $\mathbb{F}^{q_{\min}, q_{\max}}$).

Most computers use a binary floating point system, i.e., $\beta = 2$, and support multiple floating point formats. Here a “floating point format” corresponds to the choice of a set $\mathbb{F}^{q_{\min}, q_{\max}}$ for specific values of $q_{\min}$, $q_{\max}$ and $p$. The most commonly used formats are the *single* and *double* precision format, which correspond to the choices $(q_{\min}, q_{\max}, p) = (-126, 127, 24)$ and $(q_{\min}, q_{\max}, p) = (-1022, 1023, 53)$ respectively. The protocols for how these formats should be stored and the exact specifications of $q_{\min}$, $q_{\max}$ and $p$ are described in the IEEE 754 protocol, which every computer-chip manufacturer has to comply to, see [58].

For notational convenience, write $\mathbb{F}^* = \mathbb{F}^{q_{\min}, q_{\max}} \cup \{\pm \infty\}$ and $\mathbb{R}^* = \mathbb{R} \cup \{\pm \infty\}$.

To approximate elements in $\mathbb{R}^*$ computers are equipped with so-called *rounding modes*, which are order preserving mappings $\text{Rnd} : \mathbb{R}^* \rightarrow \mathbb{F}^*$ that leave elements in $\mathbb{F}^*$ fixed. In particular, all computers are equipped with the *upward* and *downward* rounding modes $\nabla, \triangle : \mathbb{R}^* \rightarrow \mathbb{F}^*$ defined by

$$
\nabla(x) := \max \{ y \in \mathbb{F}^* : y \leq x \}, \quad \triangle(x) := \min \{ y \in \mathbb{F}^* : y \geq x \}.
$$

Note that $x \in [\nabla(x), \triangle(x)]$ for any $x \in \mathbb{R}^*$. These rounding modes are used to define another rounding mode called “round to nearest”, which associates to a real number $x$ the closest floating point number in $\mathbb{F}^*$ (see Remark 1.2.1 for a more precise definition). This is the default rounding mode used on most computers.

**Remark 1.2.1.** Modern computers use a rounding mode $\text{Rnd}_2 : \mathbb{R}^* \rightarrow \mathbb{F}^*$ called round to nearest even, which is defined by

$$
\text{Rnd}_2(x) := \begin{cases} 
\nabla(x), & x < \frac{1}{2} (\nabla(x) + \triangle(x)), \\
\triangle(x), & x > \frac{1}{2} (\nabla(x) + \triangle(x)).
\end{cases}
$$

In the rare case that $x = \frac{1}{2} (\nabla(x) + \triangle(x))$, one uses the upward rounding mode if the last binary digit of $\triangle(x)$ is one, and the downward rounding mode otherwise, in order to ensure an unbiased rounding near zero (see [58] for a more thorough discussion).

A crucial shortcoming of the set $\mathbb{F}^*$ is that it is not closed under the elementary operations $\ast \in \{+,-,/,\cdot\}$. For this reason, each elementary operation $\ast$ is replaced with an approximation $\otimes : \mathbb{F}^* \times \mathbb{F}^* \rightarrow \mathbb{F}^*$. More precisely, for each rounding mode $\text{Rnd} \in \{\nabla, \triangle\}$ and operation $\ast \in \{+,-,/,\cdot\}$, there is a corresponding operation $\otimes_{\text{Rnd}}$ such that $x \otimes_{\text{Rnd}} y = \text{Rnd}(x \ast y)$ for all $x, y \in \mathbb{F}^*$. That is, the result of $x \otimes_{\text{Rnd}} y$ is the same as if we were to apply the exact operation $\ast$ to $x$ and $y$ and then used the rounding mode. The reader is referred to [58] for more details.

As mentioned before, most real numbers do not have an exact representation in $\mathbb{F}^*$. The idea of interval arithmetic is to solve this problem by replacing the floating point approximation of a real number $x$ with the interval $[\nabla(x), \triangle(x)]$ and to “lift”
the elementary operations $\star$ to intervals. To be more precise, let $\mathbb{IF}^*$ denote the set of intervals with endpoints in $\mathbb{F}^*$. Then the idea is to construct for each elementary operation $\star \in \{+, -, \cdot, /\}$ a corresponding operation $\Box : \mathbb{IF}^* \times \mathbb{IF}^* \to \mathbb{IF}^*$ such that

$$\{x_1 \star x_2 : x_1 \in I_1, \ x_2 \in I_2\} \subset I_1 \Box I_2, \ \forall I_1, I_2 \in \mathbb{IF}^*. \hspace{1cm} (1.4)$$

In particular, if $I_1, I_2 \in \mathbb{IF}^*$ are enclosures for real numbers $x, y \in \mathbb{R}^*$, respectively, then the true outcome of $x \star y$ is contained in $I_1 \Box I_2$. Hence, with the aid of interval arithmetic, one can verify inequalities on the computer and thus perform (computer-assisted) mathematical analysis.

To illustrate how simple it is to define arithmetic on $\mathbb{IF}^*$, we explain how one can define addition. Suppose $x \in I_1 = [x_1, x_2], \ y \in I_2 = [y_1, y_2]$, where $x, y \in \mathbb{R}^*$ and $I_1, I_2 \in \mathbb{IF}^*$. Then clearly

$$x + y \in [\triangledown (x_1 + y_1), \triangle (x_2 + y_2)] \in \mathbb{IF}^*.$$

Hence it is natural to define addition $\Box$ on $\mathbb{IF}^*$ by

$$[x_1, x_2] \Box [y_1, y_2] := [\triangledown (x_1 + y_1), \triangle (x_2 + y_2)].$$

Subtraction, multiplication and division can all be defined in an analogous manner.

Finally, we remark that there are many software packages and libraries which support interval arithmetic; we mention BIAS, CAPD [2] (which includes a comprehensive C++-library for validated numerical computations for both discrete and continuous-time dynamical systems) and INTLAB [66] to name a few. In particular, the validated numerical methods in this dissertation are all implemented in MATLAB using the INTLAB package, see [67].

### 1.2.2 A verification method based on Newton’s method

The validated numerical methods in this dissertation are all based on establishing the existence of a zero of a mapping $F : \mathcal{X} \to \mathcal{Y}$ by using the methodology first developed in [27, 84]. The main idea is to begin by computing an approximate zero $\hat{x}$ of a finite dimensional reduction $F_m$ of $F$ with the aid of a computer (see Section 1.1.3). Next, a mixture of analysis on paper and (floating point) computations on the computer are used to construct a Newton-like map $T$ whose fixed points correspond to zeros of $F$. Finally, the hypotheses of the contraction mapping principle are translated into a finite number of inequalities, which (if satisfied) prove that $T$ is a contraction in a neighborhood of the approximate solution $\hat{x}$. An important property of these inequalities is that they can be verified with the aid of interval arithmetic.

We stress that there exist many validated numerical methods which are based on Newton’s method. It is out of the scope of this text to give a complete overview of such methods and refer the reader to [6, 7, 15, 16, 19, 56, 63, 71, 75] for more examples.
The classical Newton method

To give a more precise description of the verification method used in this dissertation, we first review Newton’s method and explain why it is so effective. To this end, let $F : X \to Y$ be a smooth map, say at least of class $C^2$, between Banach spaces and suppose $x^*$ is a zero. Furthermore, assume that the Fréchet derivative of $F$ is invertible at $x^*$. Then there exists a radius $r > 0$ such that $F^{-1}(0) \cap B_r(x^*) = \{x^*\}$, where $B_r(x^*)$ is the closed ball of radius $r$ centered at $x^*$. In particular, we may assume without loss of generality that $DF(x)$ is invertible for all $x \in B_r(x^*)$.

Newton’s method is based on the following heuristic observation: if $x_0 \in B_r(x^*) \setminus \{x^*\}$ and $r$ is sufficiently small, then the solution of the linearized equation

$$ F(x_0) + DF(x_0)(x - x_0) = 0 $$

should be a more accurate approximation of $x^*$ than $x_0$. Hence the sequence of iterations $x_{k+1} = N(x_k)$, where $k \in \mathbb{N}_0$ and $N : B_r(x^*) \to X$ is the Newton map defined by

$$ N(x) := x - [DF(x)]^{-1} F(x), $$

is expected to converge to $x^*$ for any $x_0 \in B_r(x^*)$. This iteration scheme is referred to as Newton’s method. Note that zeros of $F$ in $B_r(x^*)$ correspond to fixed points of $N$ and vice versa. In particular, $x^*$ is the unique fixed point of $N$ in $B_r(x^*)$, since $x^*$ is the only zero of $F$ in $B_r(x^*)$.

A more precise explanation of why Newton’s method works and is so effective is that the Newton map $N$ becomes an “arbitrarily strong” contraction as one gets closer to $x^*$. An exact formulation of this statement is given in the theorem below.

**Theorem 1.2.2** (Newton’s method). Suppose $F \in C^2(X, Y)$ and $x^*$ is a zero of $F$. If $DF(x^*)$ is invertible, then for any $\varepsilon > 0$ there exists a radius $r > 0$ such that the Lipschitz constant of $N$ on $B_r(x^*)$ is strictly below $\varepsilon$. In particular, $x^*$ is a locally unique zero of $F$.

**Proof.** Let $x \in B_r(x^*)$ be arbitrary and observe that

$$ DF(x) N(x) = DF(x) x - F(x). $$

Since $N$ is continuously differentiable on $B_r(x^*)$ (because $DF$ is), we may use the Leibniz rule to differentiate both sides of the above expression, which yields

$$ D^2F(x)[h, N(x)] + DF(x) DN(x) h = D^2F(x)[h, x], \quad \forall h \in X. $$

Substitution of $x = x^*$ in the latter equation yields

$$ DF(x^*) DN(x^*) h = 0, \quad \forall h \in X, $$

since $N(x^*) = x^*$. Therefore, $DN(x^*) = 0$, since $DF(x^*)$ is injective. This proves the statement, since $N$ is $C^1$. \qed
Remark 1.2.3. The assumptions of this theorem can be relaxed. One can assume for instance that \( F \in C^1(X, Y) \) and that \( DF \) is Lipschitz (with some specific constraints). The exact conditions are stated in the Newton-Kantorovich Theorem \(^{[57]}\).

A Newton-like method

The verification methods developed in this thesis are all founded on the same basic idea. Namely, if we have numerical evidence for the existence of a (isolated non-degenerate) zero \( x^* \) of \( F \), then by Theorem 1.2.2 we should be able to rigorously prove this by establishing the contractivity of \( N \) in a small neighborhood \( B_r(\hat{x}) \) of an approximate (numerically obtained) zero \( \hat{x} \). In order for this strategy to be successful, the approximate zero \( \hat{x} \) should be sufficiently accurate and the neighborhood \( B_r(\hat{x}) \) should not be too small, since it should contain the exact zero \( x^* \).

For nonlinear problems it is usually difficult to directly prove that the Newton mapping \( N \) is a contraction. The main obstacle is the analysis of the nonlinear map \( x \mapsto [DF(x)]^{-1} \). To overcome this issue, we exploit the fact that the contraction rate of \( N \) keeps improving as we get closer to \( x^* \). More precisely, since the contraction rate gets arbitrarily strong as we approach \( x^* \), we expect that if we perturb \( N \) a little bit, the resulting map \( T \) is still a contraction (although with a weaker contraction rate) on \( B_r(\hat{x}) \) for some sufficiently small \( r > 0 \). Hence, to make the analysis feasible, we may replace the exact inverse \([DF(x)]^{-1}\) with an approximate inverse which is easier to analyze.

In practice, to get rid of the nonlinear dependence of \([DF(x)]^{-1}\) on \( x \), we use a fixed approximate inverse \( A \approx [DF(\hat{x})]^{-1} \) and define a Newton-like map \( T : X \to X \) by

\[
T := I_X - AF,
\]

where \( I_X : X \to X \) is the identity on \( X \). This idea is common in traditional numerical analysis where one typically uses approximate inverses to reduce computational costs. Note that if \( A \) is injective, then as before, fixed points of \( T \) correspond to zeros of \( F \) and vice versa.

The construction of \( A \) consists of a mixture of analysis on paper and numerical (floating point) computations on the computer. In particular, the definition of \( A \) is based on the assumption that in a neighborhood of \( x^* \) the behavior of \( F \) can be governed by a sufficiently accurate finite dimensional approximation \( F_m : X_m \to Y_m \), where \( X_m \subset X \) and \( Y_m \subset Y \) are finite dimensional subspaces (see Section 1.1.3). In fact, this assumption is the whole reason why it makes sense to perform numerical simulations in the first place. To make the discussion more precise, set

\[
\Pi_{X, \infty} := I_X - \Pi_{X,m}, \quad \Pi_{Y, \infty} := I_Y - \Pi_{Y,m},
\]

where \( I_X : X \to X \) and \( I_Y : Y \to Y \) are the identity operators on \( X \) and \( Y \), respectively. Then by using the decompositions

\[
X = \Pi_{X,m}(X) \oplus \Pi_{X,\infty}(X), \quad Y = \Pi_{Y,m}(Y) \oplus \Pi_{Y,\infty}(Y),
\]

and
we can write

\[
DF(x^*) = \begin{bmatrix}
\Pi_{\mathcal{Y},m} DF(x^*) & \Pi_{\mathcal{Y},m} DF(x^*) & \Pi_{\mathcal{Y},\infty} DF(x^*) & \Pi_{\mathcal{Y},\infty}
\end{bmatrix}.
\]

The assumption that the behavior of \( F \) near \( x^* \) is governed by \( F_m \) corresponds to the assertions that \( F_m \) and \( F - F_m \) are small near \( x^* \) and that the contribution of the operators

\[
\Pi_{\mathcal{Y},m} DF(x^*) & \Pi_{\mathcal{X},\infty} \\
\Pi_{\mathcal{Y},\infty} DF(x^*) & \Pi_{\mathcal{X},m}
\]

is negligible, i.e., the associated operator norms are relatively small.

The construction of \( A : \mathcal{Y} \to \mathcal{X} \) is based on the latter assumptions, which give rise to the approximation

\[
[DF(x^*)]^{-1} \approx \begin{bmatrix}
[DF_m(\hat{x})]^{-1} & 0 \\
0 & [\Pi_{\mathcal{Y},\infty} DF(\hat{x}) \Pi_{\mathcal{X},\infty}]^{-1}
\end{bmatrix}.
\]

In applications, we use the computer to determine an approximate inverse \( A_m \approx [DF_m(\hat{x})]^{-1} \) for the finite dimensional part of the problem. For the infinite dimensional part, we construct a linear operator \( A_\infty : \Pi_{\mathcal{Y},\infty}(\mathcal{Y}) \to \Pi_{\mathcal{X},\infty}(\mathcal{X}) \), which approximates the “dominant” part of the operator \( \Pi_{\mathcal{Y},\infty} DF(\hat{x}) \Pi_{\mathcal{X},\infty} \). The exact choice for \( A_\infty \) depends on the specific problem at hand. Finally, we set

\[
A := \begin{bmatrix}
A_m & 0 \\
0 & A_\infty
\end{bmatrix}.
\]

**Example 1.2.4.** In Example 1.1.1 a natural choice for \( A_\infty \) is

\[
[A_\infty(a)]_k := (\hat{\omega}k)^{-1} a_k, \quad |k| \geq K + 1,
\]

where \( \hat{\omega} \) is a numerical approximation of the angular frequency of the periodic orbit.

Finally, to establish that \( T \) is a contraction, we determine the quality of the approximate zero \( \hat{x} \) and compute an upper bound for the Lipschitz constant of \( T \) on \( B_r(\hat{x}) \). The quality of the approximate zero is determined by computing a bound for the residual \( \|T(\hat{x}) - \hat{x}\|_\mathcal{X} \). The Lipschitz constant is analyzed by computing a upper bound for \( \|DT(\hat{x})\|_{\mathcal{B}(\mathcal{X},\mathcal{X})} \) on \( B_r(\hat{x}) \), which involves a rigorous analysis of the error between the exact and approximate inverse at \( \hat{x} \). Since the radius \( r \) of the ball on which \( T \) is a contraction is a priori unknown, we keep it as an additional parameter in the estimates. Altogether, this information can be used to determine a neighborhood \( B_r(\hat{x}) \) on which \( T \) is a contraction. The exact conditions to be checked are stated in the theorem below, which is often referred to as a parameterized Newton-Kantorovich method.
Theorem 1.2.5 (Parameterized Newton-Kantorovich method). Let $T : X \to X$ be a Frechét differentiable mapping and $\hat{x} \in X$. Suppose there exists bounds $Y \geq 0$ and $Z : [0, \infty) \to [0, \infty)$ such that

$$\|T (\hat{x}) - \hat{x}\|_X \leq Y,$$

$$\sup_{h \in B_1(0)} \|DT (\hat{x} + rh)\|_{B(X,X)} \leq Z(r).$$

If there exists a radius $\hat{r} > 0$ such that

$$Z(\hat{r}) \hat{r} + Y < \hat{r}, \quad (1.5)$$

then $T : B_{\hat{r}} (\hat{x}) \to B_{\hat{r}} (\hat{x})$ is a contraction.

Proof. The proof is straightforward, see [84] for instance. \square

Remark 1.2.6. In the literature, this theorem is frequently referred to as the radii-polynomial approach. The reason for this terminology is that the bound $Z(r)$ can be defined as a polynomial in $r$. To see this, suppose $F \in C^M (X,Y)$ for some $M \in \mathbb{N}_{\geq 2}$. Let $h \in B_1(0)$ be arbitrary, then a Taylor-expansion shows that

$$DT (\hat{x} + rh)$$

$$= \sum_{j=0}^{M-2} \frac{r^j}{j!} D^{j+1} T (\hat{x}) h^j + r^{M-1} \int_0^1 \frac{(1-\tau)^{M-2}}{(M-2)!} D^M T (\hat{x} + \tau rh) h^{M-1} d\tau,$$

where we have used the abbreviation

$$D^{j+1} T (\hat{x}) h^j = D^{j+1} T (\hat{x}) \left[ h, \ldots, h \right]_{j \text{ times}}.$$

In applications, we choose an a-priori upper bound $r^* \geq r > 0$ for the unknown radius so that we can estimate

$$\|DT (\hat{x} + rh)\|_{B(X,X)} \leq \sum_{j=0}^{M-2} \frac{r^j}{j!} \|D^{j+1} T (\hat{x}) h^j\|_{B(X,X)}$$

$$+ \frac{r^{M-1}}{(M-1)!} \sup_{\tau \in [0,1]} \sup_{r \in (0,r^*)} \left\| D^M T (\hat{x} + \tau rh) h^{M-1} \right\|_{B(X,X)}. $$

This shows that $Z(r)$ can be defined as a polynomial in $r$. In practice, we often use $M = 2$ and construct a first order polynomial $Z(r)$ which satisfies

$$Z(r) \geq \|I_X - ADF (\hat{x})\|_{B(X,X)} + r \sup_{\tau \in [0,1]} \sup_{r \in (0,r^*)} \left\| AD^2 F (\hat{x} + \tau rh) [h, \cdot] \right\|_{B(X,X)}$$

12
for all $h \in B_1(0)$. The inequality in (1.5) can then be verified by determining an interval on which the quadratic radius polynomial $p(r) := (Z(r) - 1)r + Y$ is negative.

Remark 1.2.7. In practice, we find an interval $[r_{\text{min}}, r_{\text{max}}]$ on which (1.5) is satisfied. The radius $r_{\text{min}}$ yields information about the quality of the approximate solution. The radius $r_{\text{max}}$ yields information about the size of the largest neighborhood in which the exact solution is unique.

Theorem 1.2.5 replaces the problem of solving an equation in a Banach space with the problem of verifying a single inequality. Hence, if we construct the bounds $Y$ and $Z(r)$ in such a way that they are computable, i.e., they can be constructed with the aid of interval arithmetic, we can rigorously verify the conditions of Theorem 1.2.5 on a computer. In the end, this is what all our computer-assisted proofs boil down to: the verification of a finite number of inequalities.

1.3 Example: Chebyshev series for delay differential equations

In this section we illustrate the ideas presented in Section 1.2.2 in more detail by considering an explicit example. We consider the following initial value problem:

\[
\begin{align*}
\frac{du}{dt}(t) &= g(u(t - \tau), u(t)), \quad t \in [0, \tau], \\
u(t) &= u_0(t), \quad t \in [-\tau, 0],
\end{align*}
\]

(1.6)

where $g : \mathbb{R}^2 \to \mathbb{R}$ is a polynomial, $\tau > 0$, and $u_0 : [-\tau, 0] \to \mathbb{R}$ is a prescribed initial condition. The equation in (1.6) is referred to as a delay differential equation (DDE), or just simply as a delay equation, with constant delay $\tau$. In this section we present a computer-assisted framework for solving (1.6) with the aid of Chebyshev series. This example is complementary to the ODEs and PDEs considered in the main body of this thesis. It showcases the versatility of these Chebyshev-based techniques.

Remark 1.3.1. Delay equations are fundamentally different from ODEs. The main difference is that $u'(t)$ does not only depend on the value of $u$ at time $t$, but also on the “history” of $u$ in $[t - \tau, t]$. Therefore, in order for (1.6) to be well-posed, one has to prescribe a function on $[-\tau, 0]$ rather than just an initial value at $t = 0$. As a consequence, the phase space of the dynamical system generated by (1.6) is a function space and thus infinite dimensional. In particular, (1.6) generates a semi-flow on $C([-\tau, 0], \mathbb{R})$, see [42] for instance.

Chebyshev series We start by giving a brief introduction into the theory of Chebyshev series. The Chebyshev polynomials $T_k : [-1, 1] \to \mathbb{R}$ of the first kind are defined by the relation $T_k(\cos \theta) = \cos (k\theta)$, where $k \in \mathbb{N}_0$ and $\theta \in [0, \pi]$, ...
and constitute a non-periodic analog of the Fourier cosine basis. In particular, Chebyshev and Fourier cosine series have similar convergence properties. For the objective in this section it suffices to know that for any analytic function \( f : [-1, 1] \to \mathbb{R} \) there exist unique coefficients \( (f_k)_{k \in \mathbb{N}_0} \in \mathbb{R}^{\mathbb{N}_0} \), which decay at a geometric rate to zero, such that

\[
f = f_0 + 2 \sum_{k=1}^{\infty} f_k T_k.
\]  

The coefficients \( (f_k)_{k \in \mathbb{N}_0} \) are referred to as the Chebyshev coefficients of \( f \).

**Remark 1.3.2.** Note that (1.7) is, up to a coordinate transformation, a Fourier cosine series. This is the motivation for using the factor 2 in front of the coefficients \( (f_k)_{k \in \mathbb{N}} \). With this convention the Chebyshev coefficients of the product of two Chebyshev expansions is given directly (without a rescaling factor) by the symmetric discrete convolution.

Throughout this section we shall use the following properties of the Chebyshev polynomials:

\[
T_{k_1} T_{k_2} = \frac{1}{2} \left( T_{k_1+k_2} + T_{|k_1-k_2|} \right), \quad k_1, k_2 \in \mathbb{N}_0,
\]  

(1.8)

\[
\frac{dT_k}{dt}(t) = \frac{k}{1-T_2(t)} (T_{k-1}(t) - T_{k+1}(t)), \quad k \in \mathbb{N}, \quad t \in (-1, 1),
\]  

(1.9)

\[
T_k(-1) = (-1)^k, \quad T_k(1) = 1, \quad k \in \mathbb{N}_0.
\]  

(1.10)

These relations are direct consequences of the fact that the Chebyshev polynomials are, up to a coordinate transformation, just cosines.

**An equivalent zero finding problem** Next, we set up an equivalent zero finding problem for (1.6) by using Chebyshev series. To this end, observe that (1.6) is equivalent to

\[
\begin{cases}
\frac{d\tilde{u}}{dt}(t) = \frac{\tau}{2} g(\tilde{u}_0(t), \tilde{u}(t)), & t \in [-1, 1], \\
\tilde{u}_0(1) = \tilde{u}(-1),
\end{cases}
\]  

(1.11)

where \( \tilde{u}_0 : [-1, 1] \to \mathbb{R} \) is some prescribed analytic function. Indeed, if \( \tilde{u} \) solves (1.11), then \( u(t) := \tilde{u} \left( \frac{2t}{\tau} - 1 \right) \) solves (1.6) with initial condition

\[
u_0(t) := \tilde{u}_0 \left( \frac{2}{\tau} (t + \tau) - 1 \right).
\]

The implication in the other direction is similar. We rescaled the time domains \([-\tau, 0]\) and \([0, \tau]\) to \([-1, 1]\) for the sake of convenience, since this is the canonical domain of the Chebyshev polynomials.
Let \( a^0 = (a_k^0)_{k \in \mathbb{N}_0} \) and \( a = (a_k)_{k \in \mathbb{N}_0} \) denote the Chebyshev coefficients of \( \tilde{u}_0 \) and \( \tilde{u} \), respectively, i.e.,

\[
\tilde{u}_0 = a_0^0 + 2 \sum_{k=1}^{\infty} a_0^k T_k, \quad \tilde{u} = a_0 + 2 \sum_{k=1}^{\infty} a_k T_k. \tag{1.12}
\]

Since we are solving an initial value problem, we assume that \( a^0 \) is a-priori known.

Let \( c\left(a^0, a\right) = (c_k\left(a^0, a\right))_{k \in \mathbb{N}_0} \) denote the Chebyshev coefficients of the map \( t \mapsto g(\tilde{u}_0(t), \tilde{u}(t)) \), i.e.,

\[
g(\tilde{u}_0(t), \tilde{u}(t)) = c_0 \left(a^0, a\right) + 2 \sum_{k=1}^{\infty} c_k \left(a^0, a\right) T_k(t), \quad t \in [-1, 1].
\]

**Remark 1.3.3.** If \( \{g_\alpha : \alpha \in A\} \subset \mathbb{R} \), where \( A \subset \mathbb{N}_0^2 \), are the coefficients of \( g \) in the monomial basis, then

\[
c(b, a) = \sum_{\alpha \in A} g_\alpha b^{\alpha_1} a^{\alpha_2}, \tag{1.13}
\]

where for notational convenience we have written \( b = a^0 \). The products in (1.13) are understood to be the symmetric discrete convolution (see Chapters 2 and 3 for a more thorough discussion).

Substitution of the above expansions into (1.11) shows that

\[
\frac{d\tilde{u}}{dt} = 2 \sum_{k=1}^{\infty} a_k \frac{dT_k}{dt} = \frac{2}{1 - T_2} \left( \sum_{k=0}^{\infty} (k+1)a_{k+1}T_k - \sum_{k=2}^{\infty} (k-1)a_{k-1}T_k \right) = \tau \left( c_0 + 2 \sum_{k=1}^{\infty} c_k T_k \right), \tag{1.14}
\]

where we used (1.9) and the abbreviation \( c_k = c_k \left(a^0, a\right) \). Furthermore,

\[
(1 - T_2) \left( c_0 + 2 \sum_{k=1}^{\infty} c_k T_k \right) = (1 - T_2) c_0 - c_1 T_1 + 2 \sum_{k=1}^{\infty} c_k T_k - \sum_{k=3}^{\infty} c_{k-2} T_k - \sum_{k=0}^{\infty} c_{k+2} T_k \tag{1.15}
\]

by (1.8). Collecting terms in (1.14) and (1.15) of the same order yields the following infinite dimensional system of algebraic equations for the Chebyshev coefficients of \( \tilde{u} \):

\[
ka_k - \frac{\tau}{4} \left( c_{k-1} \left(a^0, a\right) - c_{k+1} \left(a^0, a\right) \right) = 0, \quad k \in \mathbb{N}.
\]
Moreover, the equation \( \tilde{u}_0(1) = \tilde{u}(-1) \) is equivalent to

\[
(a_0 - a_0^0) + 2 \sum_{k=1}^{\infty} \left( (-1)^k a_k - a_k^0 \right) = 0
\]

by (1.10).

We are now ready to set up an equivalent zero finding problem for (1.6). We start by defining an appropriate domain for the zero finding map \( F \). Since solutions of (1.6) are analytic, the Chebyshev coefficients of \( \tilde{u} \) decay at a geometric rate to zero. For this reason, we define the domain of \( F \) to be the space of geometrically decaying sequences with some prescribed minimal decay rate. More precisely, let \( \nu > 1 \) and set

\[
\ell_1^\nu := \{ a \in \mathbb{R}^\mathbb{N}_0 : \sum_{k=0}^{\infty} |a_k| \nu^k < \infty \}
\]

We endow this space with the norm

\[
\|a\|_\nu := |a_0| + 2 \sum_{k=1}^{\infty} |a_k| \nu^k,
\]

which turns it into a Banach space.

**Remark 1.3.4.** The space \( \ell_1^\nu \) is used extensively throughout this dissertation. The choice for the particular norm in (1.16) is twofold. First, it turns \( \ell_1^\nu \) into a Banach algebra with respect to the symmetric discrete convolution. In particular, it follows from this observation that \( c : \ell_1^\nu \times \ell_1^\nu \to \ell_1^\nu \). Secondly, the computation of operator norms with respect to a weighted \( \ell^1 \)-norm is relatively easy. We refer the reader to Chapter 3 for a more detailed and comprehensive discussion.

**Definition 1.3.5** (Chebyshev map for delay equations I). Let \( 1 < \tilde{\nu} < \nu \) and \( a^0 \in \ell_1^{\tilde{\nu}} \). The Chebyshev map \( F : \ell_1^{\tilde{\nu}} \to \ell_1^\nu \) for (1.6) is defined by

\[
(F(a))_k := \begin{cases} 
(a_0 - a_0^0) + 2 \sum_{l=1}^{\infty} \left( (-1)^l a_l - a_l^0 \right), & k = 0, \\
ka_k - \frac{\tau}{4} (c_{k-1} (a^0, a) - c_{k+1} (a^0, a)), & k \in \mathbb{N}.
\end{cases}
\]

**Remark 1.3.6.** The map \( F \) is well-defined, since \( (ka_k)_{k \in \mathbb{N}_0} \in \ell_1^\nu \) for any \( a \in \ell_1^{\tilde{\nu}} \) and \( 1 < \tilde{\nu} < \nu \), and \( c : \ell_1^{\tilde{\nu}} \times \ell_1^{\tilde{\nu}} \to \ell_1^{\tilde{\nu}} \).

**Remark 1.3.7.** Recall from Section 1.2.2 that the strategy for proving the existence of a zero of \( F \) is to construct a fixed point map from \( \ell_1^{\tilde{\nu}} \) into itself. For this reason, the weight \( \tilde{\nu} \) is irrelevant; it is only introduced to specify the codomain of \( F \).

We now have the following result by construction:

**Proposition 1.3.8.** Let \( a \in \ell_1^{\tilde{\nu}} \) and define \( \tilde{u} \) and \( \tilde{u}_0 \) as in (4.10). Then \( F(a) = 0 \) if and only if \( u(t) := \tilde{u} \left( \frac{2t}{\tau} - 1 \right) \) solves (1.6) with initial condition \( u_0(t) := \tilde{u}_0 \left( \frac{2}{\tau} (t + \tau) - 1 \right) \).
An equivalent fixed-point problem  Next, we set up an equivalent fixed-point problem for (1.6) by constructing a Newton-like map for $F$. We begin by defining an appropriate finite dimensional reduction amenable to numerical computations. To this end, let $N \in \mathbb{N}$ be a given truncation parameter and define $\Pi_N : \ell^1_\nu \to \ell^1_\nu$ by

$$(\Pi_N(a))_k := \begin{cases} a_k, & 0 \leq k \leq N - 1, \\ 0, & k \geq N. \end{cases}$$

Remark 1.3.9. In practice, we use the identification $\Pi_N(a) \simeq [a_0 \ldots a_{N-1}]^T \in \mathbb{R}^N$ and identify $\Pi_N(\ell^1_\nu) \simeq \mathbb{R}^N$.

Definition 1.3.10 (Finite dimensional reduction). The finite dimensional reduction $F_N : \mathbb{R}^N \to \mathbb{R}^N$ of $F$ is defined by

$$F_N = \Pi_N \circ F \circ \Pi_N|_{\mathbb{R}^N}.$$ 

We will construct a Newton-like map for $F$ by using a combination of numerical computations and analysis on paper. To this end, assume that we have computed an approximate zero $\hat{a} \in \mathbb{R}^N$ of $F_N$ and an approximate injective inverse $A_N$ of $DF_N(\hat{a})$ with the aid of a computer. Observe that

$$[DF(\hat{a})h]_k = kh_k - \frac{T}{4} (Dc_{k-1}(a^0, \hat{a})h - Dc_{k+1}(a^0, \hat{a})h), \quad k \in \mathbb{N}.$$ 

In particular, we expect that $[DF(\hat{a})h]_k \approx kh_k$ for $k$ sufficiently large, which motivates the following definition:

Definition 1.3.11 (Approximate inverse). The approximate inverse $A : \ell^1_\nu \to \ell^1_\nu$ of $DF(\hat{a})$ is defined by

$$Ah := \begin{bmatrix} A_N \\ N^{-1} \\ (N + 1)^{-1} \\ \vdots \end{bmatrix} \begin{bmatrix} \Pi_N(h) \\ h_N \\ h_{N+1} \\ \vdots \end{bmatrix}.$$ 

Remark 1.3.12. Observe that $A$ is injective, since $A_N$ is. Furthermore, note that $AF(a) \in \ell^1_\nu$ for any $a \in \ell^1_\nu$.

We are now ready to define a Newton-like map for $F$ based at $\hat{a}$:

Definition 1.3.13 (Newton-like map). The Newton-like map $T : \ell^1_\nu \to \ell^1_\nu$ for $F$ based at $\hat{a}$ is defined by $T := I - AF$, where $I : \ell^1_\nu \to \ell^1_\nu$ is the identity on $\ell^1_\nu$. 

17
Note that $F(x) = 0$ if and only if $T(x) = x$, since $A$ is injective. As mentioned before in Section 1.2.2, the idea is to establish the existence of a fixed point of $T$ by verifying the conditions of Theorem 1.2.5 with the aid of a computer. At this point we refer the reader to Chapter 3 where we have analyzed a very similar fixed-point problem for ODEs. The only difference between the fixed-point problem in this section and the one in Chapter 2 is the appearance of $a^0$ in the map $c$, which plays the role of a fixed parameter, not a variable. In particular, the necessary machinery for constructing the $Y$ and $Z$-bounds is developed in full detail in Chapter 3.

**Long time integration** Finally, we explain how to validate solutions of (1.6) for large integration times by solving the initial value problem

$$
\begin{aligned}
\frac{du}{dt}(t) &= g(u(t - \tau), u(t)), \quad t \in [0, \kappa \tau], \\
u(t) &= u_0(t), \quad t \in [-\tau, 0],
\end{aligned}
$$

(1.17)

where $\kappa \in \mathbb{N}$. The key observation is that (1.17) is equivalent to the system of differential equations

$$
(P_j) \begin{cases}
\frac{d\tilde{u}_j}{dt}(t) = \frac{\tau}{2} g(\tilde{u}_{j-1}(t), \tilde{u}_j(t)), & t \in [-1, 1], \\
\tilde{u}_{j-1}(1) = \tilde{u}_j(-1), & 1 \leq j \leq \kappa,
\end{cases}
$$

where $\tilde{u}_0 : [-1, 1] \to \mathbb{R}$ is some prescribed function. Indeed, if $(\tilde{u}_j)_{j=1}^\kappa$ solves $(P_j)_{j=1}^\kappa$, then a straightforward computation shows that

$$
u(t) := \sum_{j=0}^{\kappa} \tilde{u}_j \left( \frac{2}{\tau} (t - (j - 1) \tau) - 1 \right) 1_{([j-1] \tau, j \tau]}(t), \quad t \in [-\tau, \kappa \tau],
$$

(1.18)

solves (1.17). Altogether, the computations in the previous paragraph yield the following zero finding map for (1.17):

**Definition 1.3.14** (Chebyshev map for delay equations II). Let $1 < \tilde{\nu} < \nu$ and $a^0 \in \ell_\nu^1$. The Chebyshev map $F_\kappa : \prod_{j=1}^\kappa \ell_{\tilde{\nu}}^1 \to \prod_{j=1}^\kappa \ell_{\tilde{\nu}}^1$ for (1.17) is defined by

$$F_\kappa(a) := \left( f_1 \left( a^1 \right), f_2 \left( a^1, a^2 \right), \ldots, f_\kappa \left( a^{\kappa-1}, a^\kappa \right) \right),$$

where $a = (a^1, \ldots, a^\kappa)$ and $f_1 : \ell_{\tilde{\nu}}^1 \to \ell_{\tilde{\nu}}^1$, $f_j : \ell_{\tilde{\nu}}^1 \times \ell_{\tilde{\nu}}^1 \to \ell_{\tilde{\nu}}^1$ are given by

$$(f_j \left( a^{j-1}, a^j \right))_k := \left\{ \begin{array}{ll}
\left( a^j_k - a^{j-1}_k \right) + 2 \sum_{l=1}^{\infty} \left( (-1)^l a^j_l - a^{j-1}_l \right), & k = 0, \\
ka^j_k - \frac{\tau}{4} \left( c_{k-1} \left( a^{j-1}, a^j \right) - c_{k+1} \left( a^{j-1}, a^j \right) \right), & k \in \mathbb{N},
\end{array} \right.$$

for $2 \leq j \leq \kappa$. The formula for $j = 1$ is the same, but $f_1$ only depends on one variable.
As before, the following result holds by construction:

**Proposition 1.3.15.** Let \( a = (a^1, \ldots, a^\kappa) \in \prod_{j=1}^\kappa \ell^1_\nu \), set

\[
\tilde{u}_j := a_j^0 + 2 \sum_{k=1}^\infty a_j^k T_k, \quad 0 \leq j \leq \kappa,
\]

and define \( u : [-\tau, \kappa \tau] \) as in (1.18). Then \( F_\kappa(a) = 0 \) if and only if \( u \) solves (1.17).

The finite dimensional reduction, approximate inverse and Newton-like map for \( F_\kappa \) are all defined analogously by using the definitions for \( \kappa = 1 \) component wise. We refer the reader again to Chapter 3 for the details, where we have performed a very similar analysis. The main difference here is that for \( 2 \leq j \leq \kappa \) the map \( c \) depends on two variables \( (a^{j-1} \text{ and } a^j) \) instead of one; the necessary tools for dealing with this are developed in full detail in Chapter 3.

### 1.4 Outline of this thesis

The main objective in this thesis is to develop *generally applicable* validated numerical methods for the computation of invariant objects in differential equations. This includes the validated computation of periodic orbits, (un)stable manifolds and connecting orbits in nonlinear ODEs, as well as the validation of solutions of parabolic semilinear PDEs. By “generally applicable” we mean that both the mathematics and code are developed in full detail for a general class of differential equations. This is to be contrasted with the development of validated numerical methods whose sole purpose is to solve a specific problem. The two approaches are complementary. On the one hand, a general purpose method enables one to study a broad class of problems with relative ease. On the other hand, if one wishes to “push” the results for a specific problem, it is more advantageous to develop a method which exploits the “special” structure (if any) of the equation at hand.

A common theme in this dissertation is the use of *spectral methods* to analyze solutions of differential equations. In a nutshell, a spectral method corresponds to a method which is based on using an orthogonal set of basis functions, e.g. the Legendre or Fourier basis. The methods in this dissertation are all based on *Chebyshev approximations*, which constitute a non-periodic analog of Fourier cosine approximations. This particular choice is motivated by the fact that Chebyshev approximations of non-periodic functions converge at a near to optimal rate under relatively mild smoothness conditions, see [68] for instance.

We now proceed with a brief outline of the contents of this thesis. Chapters 2 and 3 are concerned with validated computations for finite dimensional dynamical systems generated by ODEs. The fourth and final chapter is concerned with validated computations for infinite dimensional dynamical systems generated by (scalar) parabolic PDEs. As mentioned before, all the verification methods in this dissertation are based on setting up an appropriate zero finding problem and a corresponding Newton-like map by using the methodology described in Section 1.2.2.
Chapter 2  In this chapter we present a computer-assisted procedure for validating analytic solutions of nonlinear ODEs by using Chebyshev series and domain decomposition. The method is developed for problems of the form

\[
\begin{aligned}
\frac{du}{dt} &= g(u, \lambda_0), \quad t \in [0, L], \\
G(u(0), u(L), \lambda_1) &= 0,
\end{aligned}
\]  

(1.19)

where \( g : \mathbb{R}^n \times \mathbb{R}^{n_0} \to \mathbb{R}^n \) is a general quadratic polynomial vectorfield in a \( n \)-dimensional phase space, which may depend on a parameter \( \lambda_0 \in \mathbb{R}^{n_0} \). The function \( G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_1} \to \mathbb{R}^{n_b} \) represents a collection of \( n_b \) boundary conditions, which may depend on a parameter \( \lambda_1 \in \mathbb{R}^{n_1} \). The main strategy is to set up (and solve) an equivalent zero finding problem for (1.19) on a space of geometrically decaying sequences by using Chebyshev series and domain decomposition. The novelty of the proposed method is the use of Chebyshev series in combination with domain decomposition. In particular, a heuristic procedure based on the theory of Chebyshev approximations for analytic functions is presented to construct efficient grids for validating solutions of (1.19). The effectiveness of the method is demonstrated by validating long periodic and connecting orbits in the Lorenz system for which validation without domain decomposition is not feasible.

While all the necessary details for proving the existence of solutions of (1.19) are presented, the main emphasis in this chapter is on the development of an efficient algorithm for performing domain decomposition. This is the reason why we restricted our attention to quadratic polynomial vectorfields. Generalization to higher order polynomial vectorfields is straightforward and is explained in detail in Chapter 3.

Chapter 3  In the third chapter we develop a general computer-assisted framework for proving the existence of transverse heteroclinic orbits connecting hyperbolic equilibria of polynomial vectorfields. To accomplish this, we consider the following system of equations:

\[
\begin{aligned}
\frac{du}{dt} &= g(u), \quad t \in [0, L], \\
u(0) &= P(\theta), \\
u(L) &= Q(\phi),
\end{aligned}
\]  

(1.20)

where

- \( g : \mathbb{R}^n \to \mathbb{R}^n \) is an arbitrary polynomial vectorfield in an \( n \)-dimensional phase space.

- \( P \) is a parameterization of a \( n_u \)-dimensional local unstable manifold of a hyperbolic equilibrium \( p_0 \in \mathbb{R}^n \). The coordinates on the associated chart are denoted by \( \theta \in \mathbb{R}^{n_u} \).
• $Q$ is a parameterization of a $n_s$-dimensional local stable manifold of a hyperbolic equilibrium $q_0 \in \mathbb{R}^n$. The coordinates on the associated chart are denoted by $\phi \in \mathbb{R}^{n_s}$.

• We assume that $n_u + n_s = n + 1$, which is a necessary condition for the existence of a tranverse connecting orbit between $p_0$ and $q_0$.

• $L > 0$ is the time of flight needed to travel between the (un)stable manifolds.

The unknowns in this problem are the orbit $u$, the parameterizations $P$ and $Q$, and the coordinates $\theta$ and $\phi$. The main idea is to compute high-order Taylor approximations of local charts on the (un)stable manifolds by using the Parameterization Method $[18,75]$ and to use Chebyshev series to parameterize the orbit in between. The existence of a connecting orbit is then characterized as a locally unique zero of a map $F$. To be more precise, the zero finding problem consists of determining the unstable and stable eigenvalues of $Dg(p_0)$ and $Dg(q_0)$, respectively, the Taylor coefficients of $P$ and $Q$, the coordinates $\theta$ and $\phi$ on the (un)stable manifolds and the Chebyshev coefficients of $u$. The novelty of the proposed method is that the (un)stable manifolds and connecting orbit in between are validated simultaneously. In particular, this approach is particularly well-suited for performing validated continuation of connecting orbits in ODEs, which is a topic of future research.

The estimates in Chapter 2 for solving $u' = g(u)$ were only developed in detail for quadratic polynomials. In this chapter we extend these estimates in full generality to deal with polynomials of any order. An additional benefit of this more general and somehow more natural viewpoint is that it resulted in somewhat sharper estimates.

The effectiveness of the method is demonstrated by validating transverse connecting orbits in four-dimensional ODEs, which arise from studying traveling waves in scalar parabolic PDEs. In particular, we prove the existence of traveling waves for a fourth order scalar parabolic PDE (an extended Fisher-Kolmogorov equation) and for a system of two scalar second order reaction-diffusion equations (Lotka-Volterra).

Chapter 4 In the fourth and final chapter we present a computer-assisted procedure for validating solutions of (scalar) semilinear parabolic PDEs, which can be extended periodically in space. We consider initial value problems of the form

\[
\begin{align*}
\frac{\partial u}{\partial t} &= \sum_{j=1}^{R} \beta_j \frac{\partial^2 u}{\partial x^2} + g(u), \quad t \in (0, L), \ x \in (0, \pi), \\
\frac{\partial u}{\partial x} (t, 0) &= \frac{\partial u}{\partial x} (t, \pi) = 0, \quad t \in [0, L], \\
u(0, x) &= f(x), \quad x \in [0, \pi].
\end{align*}
\] (1.21)

Here $g : \mathbb{R} \to \mathbb{R}$ is an arbitrary polynomial, $L > 0$ is a prescribed integration time, $f : \mathbb{R} \to \mathbb{R}$ is an analytic $2\pi$-periodic even function and $R \in \mathbb{N}$. In order to ensure
that (1.21) generates a semi-flow, we assume that $\beta_R < 0$ if $R$ is even and $\beta_R > 0$ if $R$ is odd.

The strategy is to first recast (1.21) into an equivalent infinite dimensional system of ODEs on a space of geometrically decaying sequences by using a Fourier cosine expansion in space. We then set up an equivalent zero finding problem by integrating the infinite dimensional system of ODEs with the variation of constants formula. A finite dimensional reduction of the zero finding problem is defined by approximating the dynamics of a finite number of Fourier coefficients with the aid of Chebyshev interpolation in time. A fundamental difference of the method developed in this chapter, compared to the methods developed in the previous chapters, is that the zero-finding problem is posed on a function space instead of a sequence space.

The effectiveness of the method is demonstrated by validating solutions of (1.21) for relatively large integration times for a second order PDE (Fisher’s equation) and a fourth order PDE (the Swift-Hohenberg equation).

1.5 Conclusions and future work

In Chapters 2 and 3 of this thesis we have developed generally applicable validated numerical methods for ODEs, which includes the study of periodic orbits, invariant manifolds attached to hyperbolic equilibria and connecting orbits in between. These methods are readily extendable to rigorous continuation in parameters. A topic of future research is to develop general computer-assisted methods for rigorously analyzing bifurcations of connecting orbits between hyperbolic equilibria. In addition, it would be interesting to develop similar tools for connecting orbits between other invariant sets (such as periodic orbits). Altogether, these tools will enable one to establish highly non-trivial dynamical behavior in complicated dynamical systems through forcing theorems and will aid in bridging the gap between numerical observations and mathematical proofs.

While the methods for ODEs developed in this dissertation are at a relatively advanced state, we are just at the beginning of the development of analogous tools for infinite dimensional dynamical systems, where many challenges still lie ahead. In Chapter 4 we have developed a rigorous integrator for semilinear parabolic PDEs as a first step towards the validated computation of invariant objects in infinite dimensional systems. This method is based on recasting an infinite dimensional system of ODEs into an equivalent fixed-point problem on a function space by using the variation of constants formula and Chebyshev interpolation. We believe that it would be a worthwhile project to investigate whether the method can be improved by setting up an equivalent fixed-point problem on a weighted $\ell^1$-space (instead of a function space) by using the variation of constants formula in combination with Chebyshev series (instead of Chebyshev interpolation). The reason for why this could yield an improvement is that in a weighted $\ell^1$-space one can exploit the regularity of solutions in a more direct and efficient way by prescribing the decay rates of sequences. In turn, this could aid in managing truncation errors and
facilitate the validation of orbits for large integration times.

The main motivation for developing the integrator in Chapter 4 is to provide a tool for validating saddle-to-saddle connections in semilinear parabolic PDEs. The idea is to establish the existence of a saddle-to-saddle connection by solving a finite time boundary value problem (similar to the approach in Chapter 3 for ODEs). To accomplish this goal, we need to combine the integrator with computer-assisted methods for the validated computation of local (un)stable manifolds of hyperbolic equilibria of semilinear parabolic PDEs. At this point, the only missing tool is a computer-assisted procedure for validating local stable manifolds; a computer-assisted method for validating local unstable manifolds for semilinear parabolic PDEs has already been developed in [63]. The main issue here is that stable manifolds for parabolic PDEs are infinite dimensional (as opposed to unstable manifolds which are finite dimensional). This obstructs for example the use of the parameterization method [18, 75], which requires that the stable eigenvalues of the equilibrium in question satisfy a so-called non-resonance condition. Hence, other approaches, e.g. a computer-assisted procedure based on the Lyapunov-Perron method, which consists of setting up an equivalent fixed point problem by using the variation of constants formula and an “appropriate” splitting of the phase space, need to be developed.

Finally, another interesting future research project would be to develop validated numerical methods for computing inertial manifolds for infinite dimensional dissipative systems. Inertial manifolds are smooth finite dimensional (positively) invariant manifolds, which “capture” the asymptotic behavior of a dissipative dynamical system and give a precise meaning to the statement that an infinite dimensional dynamical system exhibits finite dimensional behavior. In particular, the dynamics on an inertial manifold are governed by a finite dimensional system of ODEs. Many classical proofs for the existence of inertial manifolds are based on the Lyapunov-Perron method [35], which seems to be a good starting point for the development of computer-assisted proofs.
2 | Rigorous numerics for ODEs using Chebyshev series and domain decomposition

2.1 Introduction

In dynamical system theory one is often interested in the existence of invariant objects such as equilibria, periodic orbits, heteroclinic orbits, invariant manifolds, etc. The existence of such special orbits can reveal global information about the behavior of the dynamical system, for example through forcing theorems. The analysis of these special solutions, however, is in general difficult because of the nonlinearities in the system. Hence one usually resorts to numerical simulations. The information obtained through numerical simulation gives a lot of insight, but, unfortunately, it does not yield mathematical proofs.

The field of rigorous numerics is concerned with bridging the gap between numerical simulation and mathematically sound results. The main idea is to combine numerical simulation with analysis to establish mathematically rigorous statements. Examples of such methods can be found for instance in the CAPD software-package [2], which consists of a comprehensive C++-library for validated numerical computations of a variety of dynamically interesting objects for both discrete and continuous-time dynamical systems, using interval arithmetic Lohner-type algorithms. Another well-known software package is COSY, which is capable of rigorously integrating flows of vector fields using Taylor models [3,11].

Yet another approach is based on a parameterized Newton-Kantorovich argument, sometimes called the radii-polynomial approach. We will describe this method in full detail later in the paper. For the moment, it suffices to say that it consists of restating the problem in a fixed point formulation $T(x) = x$, and contractivity of the map $T$ (on a ball centered at the numerical approximation of the solution) is reduced to checking a finite set of inequalities that depend on the radius of the ball (i.e. the radius is a parameter), see e.g. [27,43,48,50,71].

Of particular interest for the present paper is the implementation of these ideas based on Chebyshev series introduced in [50]. Chebyshev series have, of course, long been a well-known tool in numerical analysis (see e.g. [13,53,68] and the references therein). Their successful applicability to rigorous numerics is largely
due to the analogy between Chebyshev series and Fourier series, allowing for manageable analytic estimates. The idea in [50] is to expand the unknown solution \( u \) to a boundary value problem in the Chebyshev-basis on the interval \([0, L]\), and to work, for the functional analytic arguments, on the space of algebraically decaying Chebyshev-coefficients.

The fundamental restriction of the setup in [50] is that only “short” pieces of orbit may be verified this way, since for longer orbits the coefficients in the Chebyshev series decay too slowly. Hence, although the Chebyshev series, due to their similarity to Fourier series, promise to vastly improve the efficiency of rigorous numerical algorithms (compared to, for example, spline approximations) for systems of ODEs, they thus far had the major restriction of only succeeding for short time intervals. In the current paper we solve this problem by adding domain decomposition concepts to the picture. In particular, we describe how a combination of ideas from domain decomposition and Chebyshev expansions can be united into an integrated approach for rigorous numerical computations of solutions to boundary value problems on large intervals.

To be precise, we present a rigorous numerical procedure for solving boundary value problems (BVPs)

\[
\begin{align*}
\frac{du}{dt} &= g(u, \lambda_0), & t \in [0, L], \\
G(u(0), u(L), \lambda_1) &= 0,
\end{align*}
\]

using Chebyshev series and domain decomposition. Here \( g : \mathbb{R}^n \times \mathbb{R}^{n_0} \rightarrow \mathbb{R}^n \) is a polynomial vector field in an \( n \)-dimensional phase space, which may depend on a parameter \( \lambda_0 \in \mathbb{R}^{n_0} \). We restrict our attention to polynomial vector fields for technical reasons: they allow for a relatively simple functional analytic setup, see Section 2.2, so that we can focus on the novel domain decomposition aspects. We note that many non-polynomial (but analytic) problems may be reformulated as a polynomial problem via change of variables and automatic differentiation techniques, see e.g. [48] and the references therein. Furthermore, for the sake of presentation, the estimates in Section 2.6, which are needed to validate numerical approximations of (2.1), are only developed in detail for quadratic polynomials. We remark that the estimates for higher-order polynomials are similar and straightforward generalizations of the quadratic bounds. The function \( G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_b} \) represents a collection of \( n_b \) boundary conditions, which may depend on a parameter \( \lambda_1 \in \mathbb{R}^{n_1} \).

In the BVP (2.1) parameters may either be fixed or determining their value may be part of the problem. This also holds for the length of the interval \( L \), which can be predetermined or a priori unknown (as in the case of a periodic orbit). In any case, to have a locally unique solution one needs that the number \( n_p \) of free parameters (in \( \lambda_0, \lambda_1 \) and possibly \( L \)) is such that the number of boundary conditions balances the degrees of freedom: \( n_b = n + n_p \). In the current paper we restrict our attention to such problems with locally unique solutions. We note that it is well understood how to extend the method to families of solutions via rigorous continuation techniques, see [15,27,37,73].
As explained above, the first step in the strategy extends the one presented in [50]. We recast (2.1) into an equivalent zero-finding problem in terms of the Chebyshev coefficients, where we incorporate a flexible domain decomposition component in the formulation. We then compute an approximate zero by truncation, and use a Newton-like scheme to establish the existence of the orbit of interest via a contraction argument.

The proposed method differs in one additional seminal aspect from the approach presented in [50]. The approach in [50] is based on recasting (2.1) into an equivalent fixed-point problem on the space of algebraically decaying sequences. However, integral curves of analytic vector fields are themselves analytic. Hence the associated Chebyshev coefficients decay to zero at a geometric rate rather than merely at an algebraic rate. From that perspective it is more natural to pose the equivalent fixed point problem on the space of geometrically decaying sequences, i.e., on an exponentially weighted $\ell^1$-space, see Section 2.2. This has several advantages. The estimates for bounded linear functionals and discrete convolutions, which constitute a fundamental part of the method in this paper, are more easily derived in the geometric setting; see [43] for a detailed discussion of these issues. Hence, the more transparent expressions allow us to concentrate on the core matter of domain decomposition.

Exploiting the geometric decay of the coefficients has consequences that go beyond cosmetic aspects, since the rate of decay links directly into the way the domains in the domain decomposition are chosen. We here give a brief overview of the ideas, while all details can be found in Section 2.4. We split the interval $[0, L]$ into subintervals and on each of these we write $u$ as a Chebyshev series. The main issue is how to (optimally or naturally) choose the splitting into subintervals. The theory of Chebyshev approximations explains how the decay rate of the Chebyshev coefficients of a function is related to the location of its complex singularities in the complex plane, see e.g. [68]. Crudely stated, complex singularities which are located close to the real axis are the main cause for low decay rates. A rescaling in time, or partitioning of the domain, can be used to push the complex singularities away from the real axis thereby obtaining higher decay rates (and hence fewer Chebyshev modes are needed per domain).

The goal of domain decomposition in this context is to overcome the issue of low decay rates by partitioning the domain $[0, L]$ into a finite number $m$ of subdomains, and to rigorously solve for the Chebyshev coefficients of $\{u|_{[t_{i-1}, t_i]} : 1 \leq i \leq m\}$ simultaneously. The idea is to determine a grid $\{0 = t_0 < t_1 \ldots < t_m = L\}$ such that each piece $u|_{[t_{i-1}, t_i]}$ of the orbit can be accurately approximated with a relatively small number of modes. In Section 2.4 we present a heuristic procedure for determining a grid for which the decay rates of the Chebyshev coefficients on each subdomain are (approximately) uniform over the subdomains. By choosing $m$ appropriately, this (uniform) decay is sufficiently rapid to obtain accurate approximations with a relatively small number of Chebyshev modes on each subinterval, so that a successful rigorously verified computation may ensue. The procedure is based on examining the complex singularities of the orbit to be validated by using a robust rational interpolation scheme developed in [59, 60]. In Section 2.5.2 we
illustrate that domain decomposition based on the location of the complex singularities significantly enhances the global improvement of the decay rates in a way that cannot be achieved by merely using uniform grids.

Before proceeding with a short description of some concrete results that illustrate how domain decomposition significantly enhances the applicability of Chebyshev series in computer-assisted proofs, a few remarks concerning the literature are in order. The literature on solving boundary value problems is vast, and an overview, even when restricting to rigorous computer-assisted approaches, is far beyond the scope of this paper. Let us, however, mention a few key papers to briefly sketch what kind of methods have been developed by the rigorous numerics community.

In [6,8,17,34,84] functional analytic methods, similar in spirit to ours, are used to solve BVPs: the differential equation is reformulated into an equivalent fixed-point problem and is solved by verifying the conditions of the Contraction Mapping Principle with the aid of a computer. Fundamentally different approaches based on topological rather than functional analytic methods, such as the Conley-index and covering relations, have been proven to be very effective as well (see e.g. [9,26,36,82]), especially when combined with high-accuracy interval-arithmetic integration techniques (see e.g. [10,52]). Finally, let us also mention the method in [21] based on shadowing and fixed point arguments.

To demonstrate the effectiveness of our method we have validated “long” connecting and periodic orbits in the Lorenz system

\[
\frac{du}{dt} = \begin{bmatrix}
\sigma(u_2 - u_1) \\
u_1(\rho - u_3) - u_2 \\
u_1u_2 - \beta u_3
\end{bmatrix},
\]

where \(\sigma, \beta, \rho \in \mathbb{R}\) are parameters. We set \(\beta = \frac{8}{3}, \sigma = 10\), and let \(\rho > 1\). The parameter values are referred to as classical if \(\rho = 28\). All the computations presented below have been implemented in MATLAB, using the INTLAB package [66] for the necessary interval arithmetic, and the CHEBFUN package [32] to construct the required Chebyshev approximations. The code is available at [67].

**Application 2.1.1.** We have proven the existence of a transverse heteroclinic orbit between hyperbolic equilibria in the classical Lorenz system. The implementation in [49], which in spirit is very similar to the one discussed in the present paper, except that splines are used instead of Chebyshev polynomials, was not powerful enough to verify the heteroclinic orbit for the classical parameter values. For these parameter values it turns out that the connecting orbit cannot be verified by using a Chebyshev series on a single domain, hence domain decomposition is essential.

The Lorenz system has three hyperbolic equilibria, namely the origin and \(q^\pm := (\pm \sqrt{\beta}(\rho - 1), \pm \sqrt{\beta}(\rho - 1), \rho - 1)\), which are commonly referred to as the positive
Figure 2.1: The connecting orbit from the positive eye to the origin in the Lorenz equation with classical parameters. The time of flight between the local (un)stable manifolds is $L = 30$. The geometric objects colored in red and green are representations of $W^{u}_{loc}(q^+)$ and $W^{s}_{loc}(0)$, respectively.

and negative eye. A connecting orbit from $q^+$ to the origin is characterized by

$$\begin{align*}
\frac{du}{dt} &= g(u), \quad t \in [0, L], \\
u(0) &= W^{u}_{loc}(q^+), \\
u(L) &= W^{s}_{loc}(0),
\end{align*}$$

(2.2)

where $L > 0$ is the integration time required to travel from the local unstable manifold $W^{u}_{loc}(q^+)$ of $q^+$ to the local stable manifold $W^{s}_{loc}(0)$ of the origin. The local invariant manifolds can be parameterized using the method in [49,74] to obtain rigorously validated descriptions of explicit boundary conditions that supplant the statements $u(0) \in W^{u}_{loc}(q^+)$ and $u(L) \in W^{s}_{loc}(0)$, see Section 2.5.4. The system (2.2) is thus reduced to a BVP of the type (2.1).

We established the existence of an isolated solution of (2.2) for $L = 30$ time units by using a grid consisting of $m = 55$ subdomains. The orbit is shown in Figure 2.1. The $C^0$-bound for the error between the exact and numerical approximation was of order $10^{-9}$. The reader is referred to Section 2.5.4 for the details.

Application 2.1.2. We have validated a periodic orbit of period $L \approx 25.03$ on the Lorenz attractor for the classical parameter values. We note that $L$ is a parameter and $G$ consists of $u(0) - u(L) = 0 \in \mathbb{R}^3$ plus a phase condition ($u(0)$ lies in a certain Poincaré section, see Section 2.5.2). The orbit is shown in Figure 2.2(a). The $C^0$-bound for the error between the numerical approximation and the exact solution
Figure 2.2: (a) A periodic orbit on the Lorenz attractor of period $L \approx 25.03$ validated with $m = 35$ subdomains. (b) A semi-logarithmic plot of the coefficients in the Chebyshev series on all subdomains for the three components of the solution.

was of order $10^{-10}$. Rather than pushing for extremely long orbits (which have already successfully been obtained via high-precision arithmetic [10]), this periodic orbit is primarily meant as an illustration of the typical behavior of the domain decomposition algorithm. In Figure 2.2(b) the size of the Chebyshev coefficients on all domains (as determined by the algorithm described in Section 2.4) is shown simultaneously. This showcases the fact that the algorithm determines a grid such that the decay rate is uniform.

**Application 2.1.3.** As a third application we considered a family of periodic orbits parameterized by $\rho$, accumulating to a homoclinic orbit to the origin at $\rho_{\text{hom}} \approx 13.93$. In particular, the periods of the periodic orbits tend to infinity as $\rho \downarrow \rho_{\text{hom}}$, and it becomes increasingly hard to validate the solution. Indeed, the goal of this example is to push our method to the edge of its current applicability. With the orbits spending a lot of time near the equilibrium, the algorithm for determining the domain decomposition based on the estimated location of the poles turns out to still work well for the part of the orbit that describes the near-homoclinic excursion, but not so well in the neighborhood of the equilibrium, see Section 2.5.3 for a more detailed discussion. Furthermore, the problem becomes increasingly ill-conditioned as $\rho$ approaches $\rho_{\text{hom}}$. This is remedied by considering $\rho$ as a free parameter rather than a fixed one, and adding a pseudo-arclength continuation type equation, see Section 2.5.3.

A typical validated periodic orbit near the homoclinic orbit for $\rho$ close to $\rho_{\text{hom}}$ is shown in Figure 2.3(a). It has two geometrically distinct parts: it spends a long time close to the equilibrium where the components are near-constant, see Figure 2.3(b), while the peak corresponds to the relatively short excursion into phase space. The grid is uniform on the flat part of the solution (where very few modes are used per domain), and the grid is non-uniform in the peak (where many
Figure 2.3: (a) A typical periodic orbit near the homoclinic connection. The period of the orbit is $L \approx 100.25$. Notice the sharp turn of the orbit near the origin. (b) The $x$, $y$ and $z$ components of the orbit. The three components are fairly flat for a relatively long time. These flat parts correspond to the part of the orbit near the equilibrium where the dynamics are slow.

modes are used per domain), see Section 2.5.3 for details.

This paper is organized as follows. We begin by briefly introducing the necessary background on Chebyshev series in Section 2.2. The setup for the rigorous verification of the numerical computations in the domain decomposition context is described in Section 2.3. The accompanying estimates are postponed to Section 2.6 as not to break the flow of the arguments. In Section 2.4 we discuss the algorithm for finding a domain decomposition that leads to uniform decay of the Chebyshev coefficients. Section 2.5 deals with the three applications summarized above. Finally, in Section 2.6 we first develop the full details of the estimates for the case of periodic boundary conditions, and then give the modifications required for the non-periodic boundary conditions that are used in some of the presented applications.

## 2.2 Preliminaries

### 2.2.1 Chebyshev series

The reader is referred to [68] for all proofs and a more comprehensive introduction into the theory of Chebyshev approximations. Here we summarize the properties needed for our method. The Chebyshev-polynomials $T_k : [-1, 1] \to \mathbb{R}$ of the first kind can be defined by the relation $T_k (\cos \theta) = \cos (k \theta)$, where $k \in \mathbb{N}_0$ and $\theta \in [0, \pi]$. As suggested by this definition, the Chebyshev series associated to the Chebyshev polynomials $(T_k)_{k=0}^\infty$ constitute a non-periodic analog of Fourier cosine series. In particular, Chebyshev and Fourier cosine series have similar convergence
properties. For instance, any Lipschitz continuous function admits a unique Chebyshev expansion.

The following proposition describes the decay of the Chebyshev coefficients of an analytic function.

**Proposition 2.2.1.** Suppose \( u : [-1, 1] \to \mathbb{R} \) is analytic and let

\[
 u = a_0 + 2 \sum_{k=1}^{\infty} a_k T_k
\]

be its Chebyshev expansion. Let \( E_\nu \subset \mathbb{C} \) denote an open ellipse with foci \( \pm 1 \) to which \( u \) can be analytically extended, where \( \nu > 1 \) is the sum of the semi-major and semi-minor axis. If \( u \) is bounded on \( E_\nu \), then \( |a_k| \leq M \nu^{-k} \) for all \( k \in \mathbb{N}_0 \), where \( M = \sup_{z \in E_\nu} |u(z)| \).

**Remark 2.2.2.** The largest such ellipse \( E_\nu \) is referred to as the Bernstein ellipse associated to \( u \).

The product of two Chebyshev series is (in direct analogy with Fourier series) described by a discrete convolution, as expressed by the next proposition.

**Proposition 2.2.3.** Suppose \( u, v : [-1, 1] \to \mathbb{R} \) are Lipschitz continuous and let

\[
 u = a_0 + 2 \sum_{k=1}^{\infty} a_k T_k, \quad v = b_0 + 2 \sum_{k=1}^{\infty} b_k T_k,
\]

be the associated Chebyshev expansions. Then

\[
 u \cdot v = c_0 + 2 \sum_{k=1}^{\infty} c_k T_k, \quad \text{where} \quad c = a \ast b := \sum_{k_1+k_2=k} a_{|k_1|} b_{|k_2|}.
\]

Furthermore, we state an identity which will be useful for computing the derivative of a Chebyshev-series:

\[
\frac{dT_k}{dx}(x) = \frac{k}{1 - T_2(x)} (T_{k-1}(x) - T_{k+1}(x)), \quad k \in \mathbb{N}. \tag{2.3}
\]

Finally, we have the product formula

\[
 T_{k_1} T_{k_2} = \frac{1}{2} (T_{k_1+k_2} + T_{|k_1-k_2|}). \tag{2.4}
\]

As we will be solving boundary value problems, we observe that \( T_k(1) = 1 \) and \( T_k(-1) = (-1)^k \).
2.2.2 Sequence spaces

The functional analytic reformulation of (2.1) in terms of the Chebyshev coefficients is posed on a weighted $\ell^1$ space. More precisely, in light of Proposition 3.2.2, we define the space

$$\ell^1_{(\nu,n)} := \left\{ (a_k)_{k \in \mathbb{N}_0} : a_k \in \mathbb{R}^n, \left| a_0 \right| + 2 \sum_{k=1}^{\infty} \left| a_k \right| \nu^k < \infty, 1 \leq j \leq n \right\},$$

where $[a_k]_j$ denotes the $j$-th component of $a_k \in \mathbb{R}^n$ and $\nu > 1$ is a given weight, endowed with the norm

$$\|a\|_{(\nu,n)} := \max_{1 \leq j \leq n} \left\{ \left| a_0 \right| + 2 \sum_{k=1}^{\infty} \left| a_k \right| \nu^k \right\}.$$

We shall write $\ell^1_{\nu} := \ell^1_{(\nu,1)}$ and $\|\cdot\|_\nu = \|\cdot\|_{(\nu,1)}$. The convolution $a \ast b$ of two vector-valued sequences $a, b \in \ell^1_{(\nu,n)}$ is defined component-wise.

A particularly important reason for choosing the above norm is that it induces a natural Banach algebra structure on $\ell^1_{(\nu,n)}$ with respect to the discrete convolution:

**Proposition 2.2.4.** The space $\left( \ell^1_{(\nu,n)}, \ast \right)$ is a Banach algebra. In particular

$$\|a \ast b\|_{(\nu,n)} \leq \|a\|_{(\nu,n)} \|b\|_{(\nu,n)}$$

for any $a, b \in \ell^1_{(\nu,n)}$.

Finally, we state an elementary result about the dual of $\ell^1_{\nu}$ which will be used extensively throughout this paper. Let $\{\varepsilon_l\}_{l=0}^{\infty}$ be the set of “corner points” of the unit one ball in $\ell^1_{\nu}$:

$$(\varepsilon_0)_k := \begin{cases} 1 & k = 0 \\ 0 & k > 0 \end{cases} \quad \text{and} \quad (\varepsilon_l)_k := \begin{cases} \frac{1}{2\nu^l} & k = l \\ 0 & k \neq l \end{cases} \quad \text{for } l \in \mathbb{N}. \quad (2.5)$$

Then we have the following characterization of the dual of $\ell^1_{\nu}$:

**Lemma 2.2.5.** Let $\psi \in (\ell^1_{\nu})^*$, then

$$\|\psi\|_{\nu}^* = \sup_{l \in \mathbb{N}_0} |\psi(\varepsilon_l)|.$$

**Proof.** See Proposition 3.2.8. \qed

2.3 Rigorous numerics for periodic orbits

In this section we introduce a rigorous numerical method for solving a special case of (2.1), namely we consider the problem of validating a periodic orbit. The reason
why we have chosen to present the details of the method for periodic orbits is only for the sake of clarity, and it will be shown in Section 2.5 how one can adapt the procedure to deal with other types of BVPs.

Since periodic orbits are invariant under translations in time, we need to introduce an additional phase condition in order to isolate the periodic orbit of interest. Note that a periodic orbit can be characterized by the following BVP:

\[
\begin{cases}
\frac{du}{dt} = \omega^{-1} g(u), & t \in [0, 1], \\
u(0) = u(1), \\
\langle v_0, u_0 - u(0) \rangle = 0
\end{cases}
\]  

(2.6)

where \(\omega > 0\) is the frequency of \(u\), and \(u_0, v_0 \in \mathbb{R}^n\) are fixed. The vectors \(u_0\) and \(v_0\) define a Poincaré section through which the periodic orbit \(u\) is required to pass at time \(t = 0\) (i.e. the phase condition). Note that the frequency \(\omega\) is a-priori unknown and must be included as an additional variable to solve for.

We start by recasting the problem into an equivalent zero finding problem of the form \(F(x) = 0\) in terms of the Chebyshev coefficients. Next, we construct a Newton-like operator \(T\) for \(F\) based at an approximate zero \(\hat{x}\) obtained via numerical simulation. Finally, we use a parameterized Newton-Kantorovich method to determine a finite number of explicit inequalities, which can be rigorously verified with the aid of a computer, in order to establish that \(T\) is a contraction on a neighborhood of the approximate solution \(\hat{x}\).

2.3.1 Chebyshev operator for periodic orbits

In this section we reformulate (2.6) as an equivalent equation of the form \(F(x) = 0\) by performing domain decomposition and using a Chebyshev expansion on each subdomain. Let

\[\mathcal{P}_m := \{t_0 = 0 < t_1 < \ldots < t_m = 1\}\]

be any partition of \([0, 1]\), where \(m \in \mathbb{N}\) is the mesh size. Then (2.6) is equivalent to

\[
\begin{cases}
\frac{du_1}{dt} = \omega^{-1} g(u_1), & t \in [0, t_1], \\
u_1(0) - u_m(1) = 0, \\
\langle v_0, u_0 - u_1(0) \rangle = 0,
\end{cases}
\]

\((P_1)\)

\[
\begin{cases}
\frac{du_i}{dt} = \omega^{-1} g(u_i), & t \in [t_{i-1}, t_i], \\
u_i(t_{i-1}) = u_{i-1}(t_{i-1}),
\end{cases}
\]

\((P_i)\)
where $2 \leq i \leq m$. Note that each $u_i$ (if it exists) is analytic, since $g$ is assumed to be polynomial (say of degree $N_g$):

$$g(u) = \sum_{|\alpha|=0}^{N_g} g_{\alpha} u^{\alpha},$$

where $g_{\alpha} \in \mathbb{R}^n$. Here $\alpha = (\alpha_1, \ldots, \alpha_n)$ is the usual multi-index, with $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Therefore, the Chebyshev expansion

$$u_i = a^i_0 + 2 \sum_{k=1}^{\infty} a^i_k T^i_k, \quad a^i_k \in \mathbb{R}^n,$$

(2.7)

is unique and converges uniformly to $u_i$ on $[t_{i-1}, t_i]$. Furthermore, the coefficients

$$\left(\left[a^i_k\right]_j\right)_{k \in \mathbb{N}_0}$$

where $1 \leq j \leq n$, decay geometrically to 0 as $k \to \infty$ by Proposition 3.2.2. In particular, there exist numbers $(\nu_i)_i^{m}$, where each $\nu_i > 1$, such that $\left[a^i\right]_j \in \ell^1_{\nu_i}$ for all $1 \leq i \leq m, 1 \leq j \leq n$. In the remainder of this section the weights $\nu = (\nu_i)_i^{m}$ are assumed to be fixed.

To obtain a reformulation of $(P_i)_{i=1}^{m}$ in terms of the coefficients $a^i$, first observe

$$g \circ u_i = c^i_0 + 2 \sum_{k=1}^{\infty} c^i_k T^i_k,$$

where

$$c^i = \sum_{|\alpha|=0}^{N_g} g_{\alpha} \left[a^i\right]_1^{\alpha_1} \cdots \left[a^i\right]_n^{\alpha_n},$$

(2.8)

$1 \leq i \leq m, 1 \leq j \leq n$, by Proposition 3.2.3. Note that $c^i$ is a function of $a^i$. In particular, $c^i : \ell^1_{(\nu_i,n)} \to \ell^1_{(\nu_i,n)}$, since

$$\left\| [c^i]_j \right\|_{\nu_i} \leq \sum_{|\alpha|=0}^{N_g} |g_{\alpha}|_j \prod_{l=1}^{n} \left\| \left[a^i\right]_l \right\|^{\alpha_l}_{\nu_i} < \infty,$$

(2.9)

for all $1 \leq j \leq n$, by Proposition 2.2.4. We shall write $c^i = c^i(a^i)$ whenever we need to emphasize this dependency in a more explicit way.

Substitution of the Chebyshev expansion of $u_i$ into the differential equation in $(P_i)$ yields

$$\frac{du_i}{dt} = 2 \sum_{k=1}^{\infty} a^i_k \frac{dT^i_k}{dt} = g \circ u_i = c^i_0 + 2 \sum_{k=1}^{\infty} c^i_k T^i_k.$$

(2.10)

By differentiating the Chebyshev polynomials, equating coefficients of the same order on the left- and right-hand side of (2.10), and using (2.3) and (2.4), or more
directly using $T_k(\cos \theta) = \cos(k\theta)$, one obtains an equivalent formulation of the
differential equation in terms of the coefficients $(a^1_k)_{k \in \mathbb{N}_0}$:

$$\omega ka^1_k = \frac{t_{i+1} - t_i}{4}(c^i_{k-1} - c^i_{k+1}).$$

The equivalent equations for the boundary conditions are obtained in a similar
fashion. In particular, substitution of the Chebyshev expansion of $u_i$ into the phase
condition yields

$$\langle v_0, u_0 \rangle - \langle v_0, a^1_0 \rangle - 2 \sum_{k=1}^{N_1} (-1)^k \langle v_0, a^1_k \rangle = 0,$$

where we have, without loss of generality, adapted it to depend only on finitely
many coefficients (this simplifies the estimates). In practice, $N_1$ will be the number
of modes up to which $a^1$ is computed numerically.

We are now ready to define the desired map $F$:

**Definition 2.3.1** (Chebyshev operator for periodic orbits). Let $\nu = (\nu_i)_{i=1}^m$ and
$\tilde{\nu} = (\tilde{\nu}_i)_{i=1}^m$ be collections of weights such that $1 < \tilde{\nu}_i < \nu_i$ for all $1 \leq i \leq m$.
The Chebyshev operator for periodic orbits is the map $F : \mathbb{R} \times \prod_{i=1}^m \ell^1_1(\nu_i, n) \rightarrow \mathbb{R} \times \prod_{i=1}^m \ell^1_1(\tilde{\nu}_i, n)$ defined by

$$F(x) := (f_0(a^1), f_1(\omega, a^1, a^m), f_2(\omega, a^1, a^2), \ldots, f_m(\omega, a^{m-1}, a^m)),$$

where $x = (\omega, a^1, a^m)$, $f_0 : \ell^1_1(\nu_i, n) \rightarrow \mathbb{R}$, and $f_i : \mathbb{R} \times \ell^1_1(\nu_{i-1}, n) \times \ell^1_1(\nu_i, n) \rightarrow \ell^1_1(\tilde{\nu}_i, n)$,
are given by

$$f_0(a^1) := \langle v_0, u_0 \rangle - \langle v_0, a^1_0 \rangle - 2 \sum_{k=1}^{N_1} (-1)^k \langle v_0, a^1_k \rangle,$$

$$f_i(\omega, a^{i-1}, a^i) := \begin{cases} a^i_0 - a^{i-1}_0 + 2 \sum_{k=1}^{\infty} \left((-1)^k a^i_k - a^{i-1}_k\right), & k = 0, \\ \omega ka^1_k - t_i - t_{i-1} - 4 \left(c^i_{k-1} - c^i_{k+1}\right), & k \in \mathbb{N}, \end{cases}$$

for $i = 1, \ldots, m$, where we set $a^0 = a^m$.

**Remark 2.3.2.** If $a^i \in \ell^1_1(\nu_i, n)$, then $(ka^1_k)_{k \in \mathbb{N}_0} \in \ell^1_1(\tilde{\nu}_i, n)$ for any $1 < \tilde{\nu}_i < \nu_i$.

By construction, we now have the following:

**Proposition 2.3.3.** $F(x) = 0$ if and only if the functions

$$\left\{ u_i = a^1_0 + 2 \sum_{k=1}^{\infty} a^1_k T^i_k : 1 \leq i \leq m \right\}$$

constitute a periodic orbit of $g$. 

36
2.3.2 Finite Dimensional Reduction

In this section we explain how to discretize the equation \( F(x) = 0 \) in order to compute a finite-dimensional approximate solution of (2.6). We start by introducing some notation: define the space \( X_\nu := \mathbb{R} \times \prod_{i=1}^{m} \ell_{1}(\nu_i,n) \) endowed with the norm

\[
\| (\omega, a^{1}, \ldots, a^{m}) \|_{X_\nu} := \max \{ |\omega|, \max \{ \| a^{i} \|_{(\nu_i,n)} : 1 \leq i \leq m \} \}.
\]

Define projections \( \Pi_0 : X_\nu \to \mathbb{R} \), \( \Pi_i : X_\nu \to \ell_{1}(\nu_i,n) \) and \( \Pi_{i,j} : X_\nu \to \ell_{1}(\nu_i,n) \) by

\[
\Pi_0 (\omega, a^{1}, \ldots, a^{m}) := \omega, \quad \Pi_i (\omega, a^{1}, \ldots, a^{m}) := a^{i},
\]

\[
\Pi_{i,j} (\omega, a^{1}, \ldots, a^{m}) := \begin{bmatrix} a^{i} \end{bmatrix}_j,
\]

where \( 1 \leq i \leq m \), \( 1 \leq j \leq n \).

Let \( N = (N_1, \ldots, N_m) \in \mathbb{N}^m \) be a given collection of truncation parameters and define \( \Pi_{N_i} : \ell_{1}(\nu_i,n) \to \ell_{1}(\nu_i,n) \), where \( 1 \leq i \leq m \), by

\[
\Pi_{N_i} (a^{i}) := \begin{cases} a^{i}_k, & 0 \leq k \leq N_i - 1, \\ 0_n, & k \geq N_i, \end{cases}
\]

and the Galerkin projection \( \Pi_N : X_\nu \to X_\nu \) by

\[
\Pi_N (\omega, a^{1}, \ldots, a^{m}) := (\omega, \Pi_{N_1} (a^{1}), \ldots, \Pi_{N_m} (a^{m})).
\]

Henceforth we shall identify \( \Pi_N (\omega, a^{1}, \ldots, a^{m}) \) and \( \Pi_{N_i} (a^{i}) \) with

\[
(\omega, (a^{1}_k)_{k=0}^{N_1-1}, \ldots, (a^{m}_k)_{k=0}^{N_m-1}) \in \mathbb{R}^{1+n \sum_{i=1}^{m} N_i} \quad \text{and} \quad (a^{i}_0, \ldots, a^{i}_{N_i-1}) \in \mathbb{R}^{n N_i},
\]

respectively. This is a slight abuse of notation, but it reduces clutter. It should be clear from the context when to interpret a variable in the finite dimensional space and when to interpret it as an element of an infinite dimensional space with zeros in the tail. Finally, set \( X_\nu^N := \Pi_N (X_\nu) \simeq \mathbb{R}^{1+n \sum_{i=1}^{m} N_i} \).

Definition 2.3.4 (Finite dimensional reduction of \( F \)). The finite dimensional reduction of \( F \) is the map \( F_N : X_\nu^N \to X_\nu^N \) defined by

\[
F_N \left( [\omega, (a^{1}_k)_{k=0}^{N_1-1}, \ldots, (a^{m}_k)_{k=0}^{N_m-1}] \right) := \Pi_N \left( F (\omega, a^{1}, \ldots, a^{m}) \right).
\]

2.3.3 A Newton-like scheme

In this section we introduce a method for proving the existence of an exact zero of \( F \) by using an approximate zero of \( F_N \). The main idea is to build a Newton-like scheme in the infinite dimensional setting by using approximate data obtained via numerical simulation. Assume that we have computed the following:
(C1) An approximate zero \( \hat{x} \in \mathbb{R}^{1+n \sum_{i=1}^{m} N_i} \) of \( F_N \), where \( \hat{x} = (\hat{\omega}, \hat{a}^1, \ldots, \hat{a}^m) \) and \( \hat{\omega} > 0 \).

(C2) An approximate injective inverse \( A_N \) of \( DF_N(\hat{x}) \).

The finite dimensional data will be used to construct a Newton-like operator \( T \) for \( F \) such that the zeros of \( F \) will correspond to fixed points of \( T \) and vice versa.

We start by constructing approximations of \( DF(\hat{x}) \) and its inverse by extending \( DF_N(\hat{x}) \) and \( A_N \) to \( \mathcal{X}_T \) and \( \mathcal{X}_P \), respectively. Recall that \( c^i(a^i) \) decays geometrically to \( 0 \) as \( k \to \infty \) by (2.9), for any \( a^i \in \ell_{(\nu, m)}^1 \). Moreover, \( c^i_k(\hat{a}^i_k) = 0 \) for all \( k > N_g(N_i - 1) \) by (2.8), since \( \hat{a}^i_k = 0 \) for all \( k \geq N_i \). Therefore, if the truncation sizes \( N_i \) are sufficiently large, and \( \max \{t_i - t_{i-1} : 1 \leq i \leq m \} \) is sufficiently small, the linear part of \( F \) corresponding to \( \omega k a^i_k \), where \( k \geq N_i \), will be dominant at \( \hat{x} \). Consequently, one can construct approximations of \( DF(\hat{x}) \) and its inverse by using \( DF_N(\hat{x}) \) and \( A_N \) for the finite dimensional part, respectively, and the linear part of the tail of \( F \) for the remainder:

**Definition 2.3.5** (Approximation of \( DF(\hat{x}) \)). The approximate derivative \( \hat{A} : \mathcal{X}_T \to \mathcal{X}_P \) of \( F \) at \( \hat{x} \) is defined by

\[
\Pi_0 \hat{A}(x) := \Pi_0 DF_N(\hat{x}) \Pi_N(x),
\]

\[
(\Pi_i \hat{A}(x))_k := \begin{cases} (\Pi_i DF_N(\hat{x}) \Pi_N(x))_k, & 0 \leq k \leq N_i - 1, \\ \hat{\omega} k (\Pi_i(x))_k, & k \geq N_i, \end{cases}
\]

where \( 1 \leq i \leq m \).

**Definition 2.3.6** (Approximate inverse of \( DF(\hat{x}) \)). The approximate inverse \( A \) of \( DF(\hat{x}) \) on \( \mathcal{X}_P \) is defined by

\[
\Pi_0 A(x) := \Pi_0 A_N \Pi_N(x),
\]

\[
(\Pi_i A(x))_k := \begin{cases} (\Pi_i A_N \Pi_N(x))_k, & 0 \leq k \leq N_i - 1, \\ \frac{1}{\hat{\omega} k} (\Pi_i(x))_k, & k \geq N_i. \end{cases}
\]

where \( 1 \leq i \leq m \).

**Remark 2.3.7.** The operator \( A \) is injective: suppose \( Ax = 0 \), then \( \Pi_N(x) = 0 \), since \( A_N \) is assumed, see (C2) above, to be injective, and \((\Pi_i(x))_k = 0 \in \mathbb{R}^n \) for all \( 1 \leq i \leq m \) and \( k \geq N_i \), i.e. \( x = 0 \).

In analogy to the classical notion of a Newton-operator for a finite dimensional map, we now define an infinite dimensional Newton-like operator for \( F \), based at \( \hat{x} \), as follows:

**Definition 2.3.8** (Newton-like operator for \( F \)). The Newton-like operator \( T : \mathcal{X}_T \to \mathcal{X}_T \) for \( F \), based at \( \hat{x} \), is defined by

\[
T(x) := x - AF(x).
\]
An immediate consequence of the fact that $A$ is injective is the following:

**Proposition 2.3.9.** $T(x) = x$ if and only if $F(x) = 0$.

If $\hat{x}$ is a sufficiently accurate approximate zero of $F$, we expect to find an *exact* zero $x^*$ of $F$, i.e., a fixed point of $T$, in a small neighborhood of $\hat{x}$. Moreover, if $r > 0$ is sufficiently small (not too small), we anticipate $T$ to be a contraction on $B_r(\hat{x})$. To see why, let $x_1, x_2 \in B_r(0)$, $r > 0$ be arbitrary and consider the following factorization:

$$DT(\hat{x} + x_1)x_2 = (I - ADF(\hat{x} + x_1))x_2 = (I - A\hat{A})x_2 - A\left(DF(\hat{x} + x_1) - \hat{A}\right)x_2. \quad (2.11)$$

The first term in (2.11) is related to the numerical part of the problem and measures the quality of the approximate inverse $A_N$, since $I - A\hat{A}$ vanishes in the tail while the finite dimensional part is the matrix $I_N - A_NDF_N(\hat{x})$. In particular, it is expected to be small by construction. The second term is of a more fundamental nature and involves the analysis of the infinite dimensional operator $DF$ in a neighborhood of the numerical approximation $\hat{x}$. Intuitively, we expect the difference $(DF(\hat{x} + x_1) - \hat{A})x_2$ to be small for small $x_1$ if $\hat{A}$ is an accurate approximation of $DF$ near $\hat{x}$. As mentioned before, this is likely to be true if the truncation sizes $N_i$ are sufficiently large, and the mesh-size and radius $r$ are sufficiently small. Altogether, these observations explain why it is plausible for $T$ to be contracting near $\hat{x}$.

The following theorem quantifies the above assertions and is amenable to rigorous numerical analysis. The proof is the same as for the case of a single domain, see [27].

**Theorem 2.3.10** (Contraction mapping principle with variable radius). *Assume that the following conditions are satisfied:

(i) There exist bounds $Y_{i,j}$, $Z_{i,j}(r) \geq 0$ such that

$$||\Pi_{i,j}(T(\hat{x}) - \hat{x})||_{\nu_i} \leq Y_{i,j}, \quad (2.12)$$

$$\sup_{x_1, x_2 \in B_r(0)} ||\Pi_{i,j}DT(\hat{x} + x_1)x_2||_{\nu_i} \leq Z_{i,j}(r), \quad (2.13)$$

for all $1 \leq i \leq m$, $1 \leq j \leq n$, and bounds $Y_0$, $Z_0(r) \geq 0$ such that

$$|\Pi_0(T(\hat{x}) - \hat{x})| \leq Y_0, \quad (2.14)$$

$$\sup_{x_1, x_2 \in B_r(0)} |\Pi_0 DT(\hat{x} + x_1)x_2| \leq Z_0(r), \quad (2.15)$$

where $r > 0$.\"
(ii) There exists a radius $\hat{r} > 0$ such that

$$\max \left\{ \max_{1 \leq i \leq m, 1 \leq j \leq n} \{ Z_{i,j}(\hat{r}) + Y_{i,j} \}, \ Z_0(\hat{r}) + Y_0 \right\} < \hat{r}.$$  

Then $T : B_{\hat{r}}(\hat{x}) \to B_{\hat{r}}(\hat{x})$ is a contraction.

**Remark 2.3.11.** The $Y$-bounds measure the accuracy of the approximate solution $\hat{x}$, while the $Z$-bounds measure the contraction rate of the Newton-like operator $T$.

The $Z$-bounds are polynomials in $r$, as will be shown in Section 2.6.2 which motivates the following terminology:

**Definition 2.3.12** (Radii-polynomials). The radii-polynomials for $T$ are defined by

$$p_{i,j}(r) := Z_{i,j}(r) + Y_{i,j} - r, \quad p_0(r) := Z_0(r) + Y_0 - r,$$

where $1 \leq i \leq m$, $1 \leq j \leq n$.

**Corollary 2.3.13.** If $p_0(\hat{r}), p_{i,j}(\hat{r}) < 0$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$, where $\hat{r} > 0$, then $T : B_{\hat{r}}(\hat{x}) \to B_{\hat{r}}(\hat{x})$ is a contraction.

Note that Corollary 2.3.13 also provides a rigorous error-bound for the approximate solution:

**Proposition 2.3.14.** Suppose $x^*$ is the fixed point of $T$ in $B_{\hat{r}}(\hat{x})$, then

$$\| u^* - \hat{u} \|_\infty \leq \hat{r}, \quad | \omega^* - \hat{\omega} | \leq \hat{r},$$

where $u^*, \hat{u} : [0,1] \to \mathbb{R}^n$ are the exact and approximate periodic orbit with frequency $\omega^*$ and $\hat{\omega}$ defined by $x^*$ and $\hat{x}$, respectively.

### 2.4 Domain decomposition

In this section we present a procedure, partially based on heuristics, for computing an efficient grid $\mathcal{P}_m$ which facilitates the rigorous validation process. The main idea is to compute a grid for which the decay rates of the coefficients $\hat{a}^i$ are sufficiently high and uniformly distributed over the subdomains. The motivation for this choice is based on the observation that a combination of high-decay rates (uniformly distributed) and a relatively small number of modes will help to control the tail estimates in Lemma 2.6.10 on each subdomain in a uniform way. In turn, this will aid in verifying that $T$ is a contraction.
2.4.1 A heuristic procedure for computing $\mathcal{P}_m$

In this section we introduce a heuristic procedure for computing a grid $\mathcal{P}_m$ such that the decay rate of the Chebyshev coefficients $[\hat{a}^t]_j$, where $1 \leq j \leq n$, is the same on each subdomain. The main idea is to construct $\mathcal{P}_m$ by using the Bernstein ellipses introduced in Proposition 3.2.2. Suppose $u^* : [0, 1] \to \mathbb{R}^n$ is the exact solution of the ODE under consideration (assuming it exists). Furthermore, assume that the obstructions for analytically extending the components $[u^*_j]$ to the entire complex plane are the presence of poles $\{z_{k,j}\}_{k=1}^{N_{p,j}}$.

Write $u^*_i := u^*|_{[t_{i-1}, t_i]}$ and observe that the Bernstein ellipse associated to the map

$$t \mapsto \left[ u^*_i \left( \frac{t_i - t_{i-1}}{2} (t + 1) + t_{i-1} \right) \right]_j,$$

where $t \in [-1, 1]$, i.e., the largest ellipse with foci $-1$ and $1$ to which the latter map can be analytically extended, has the following measurements: its linear eccentricity is equal to 1, the length of the semi-major axis is equal to

$$P_j (t_{i-1}, t_i) := \min_{1 \leq k \leq N_{p,j}} \left| z_{k,j} - t_{i-1} \right| + \left| z_{k,j} - t_i \right| \left| t_i - t_{i-1} \right|,$$

and the length of the semi-minor axis is equal to $\sqrt{P_j (t_{i-1}, t_i)^2 - 1}$. Therefore, the decay rate of the Chebyshev coefficients of $[u^*_i]_j$ is given by

$$P_j (t_{i-1}, t_i) + \sqrt{P_j (t_{i-1}, t_i)^2 - 1},$$

and the length of the semi-minor axis is equal to $\sqrt{P_j (t_{i-1}, t_i)^2 - 1}$. Therefore, the decay rate of the Chebyshev coefficients of $[u^*_i]_j$ is given by

$$P_j (t_{i-1}, t_i) + \sqrt{P_j (t_{i-1}, t_i)^2 - 1},$$

due to Proposition 3.2.2. We shall refer to the latter quantity as the size of the Bernstein ellipse of $[u^*_i]_j$. Note that if the components of the vector field are all coupled, the components of $u^*$ will (generically) have the same poles. In this case the decay rate will be the same for all $1 \leq j \leq n$.

Set $P (t_i, t_{i-1}) := \min_{1 \leq j \leq n} P_j (t_{i-1}, t_i)$ and note that the smallest Bernstein ellipse of the components of $u^*_i$ has size

$$\nu_e := P (t_{i-1}, t_i) + \sqrt{P (t_{i-1}, t_i)^2 - 1}.$$

Hence, equidistributing decay rates of $\hat{a}^t$ corresponds to equidistributing $P(t_{i-1}, t_i)$.

Next, define $\Phi_m : \mathcal{G}_m \subset \mathbb{R}^{m-1} \to \mathbb{R}^{m-1}$ by

$$\Phi_m (t_1, \ldots, t_{m-1}) := \begin{pmatrix} P (t_0, t_1) - P (t_1, t_2) \\ \vdots \\ P (t_{m-2}, t_{m-1}) - P (t_{m-1}, t_m) \end{pmatrix},$$

where $\mathcal{G}_m := \{(t_1, \ldots, t_{m-1}) \in \mathbb{R}^{m-1} : 0 < t_1 < \ldots < t_{m-1} < 1\}$, and observe that the zeros of $\Phi_m$ characterize the grids which equidistribute (2.19) over the subdomains $[t_{i-1}, t_i]$. Therefore, the desired grid can be obtained by computing a zero of $\Phi_m$.  

41
We shall approximate a zero of $\Phi_m$ by using Newton’s method. In order for Newton’s method to be successful, however, we need to supply a sufficiently accurate initial guess for a zero of $\Phi_m$. To find such an initial guess, we interpret $\Phi_m$ as a smooth vector field on $\mathcal{G}_m$, and the desired grid as a steady state of the associated dynamical system. This interpretation makes sense, since $\mathcal{G}_m$ is invariant under the flow induced by $\Phi_m$, i.e., the ordering of the grid points $0 < t_1 < \ldots < t_{m-1} < 1$ is preserved under the flow. The reason for this is that successive grid points $t_i - 1$ repel each other whenever their mutual distance is sufficiently small due to the factors $\pm \frac{1}{t_i - t_{i-1}}$ in $[\Phi_m(t_1, \ldots, t_{m-1})]_{i-1}$ and $[\Phi_m(t_1, \ldots, t_{m-1})]_i$, see (2.17).

If $\mathcal{G}_m$ contains a stable equilibrium of $\Phi_m$, one can approximate its location, i.e., compute an initial guess for the desired grid, by integrating the ODE
\[
\frac{dt}{d\tau} = \Phi_m(t), \quad \tau \in [0, \tau_0],
\]
for sufficiently large $\tau_0 > 0$. In practice, we start with a uniformly distributed grid and follow the flow of (2.20) for some time. In all our numerical experiments this process appeared to converge to an equilibrium state and yielded a sufficiently accurate initial guess for initiating Newton’s method.

### 2.4.2 Approximation of the complex singularities

In the previous section we explained how to compute grids by computing zeros of $\Phi_m$. In order to construct the map $\Phi_m$, however, one needs to determine the complex singularities of the exact solution $u^*$. In this section we outline some of the algorithmic aspects for approximating the relevant complex singularities of $u^*$, i.e., the ones which determine the sizes of the Bernstein ellipses, by using an approximate solution of the ODE and the rational interpolation scheme developed in [59,60].

The main idea in [59,60] is as follows: given an analytic function $f : [a, b] \to \mathbb{R}$ approximate its analytic extension into the complex plane by constructing a rational interpolant $\frac{p}{q}$. This is accomplished by sampling $f$ at the Chebyshev points in $[a, b]$ and solving (if necessary in a least square sense) the problem
\[
p(y_j) - f(y_j)q(y_j) = 0, \quad 0 \leq j \leq K,
\]
where $(y_j)_{j=0}^K$ are the Chebyshev points on $[a, b]$, $K \in \mathbb{N}$, and $p, q$ are polynomials of degree $N_p, N_q$, respectively. This will yield a rational interpolant $\frac{p}{q}$, provided $q(y_j) \neq 0$ for all $0 \leq j \leq K$. The associated rational interpolant is referred to as a rational interpolant of type $(N_p, N_q)$. The complex singularities of $f$ can be approximated by computing the roots of $q$. The degrees of $p$ and $q$, however, should be chosen carefully in order to avoid spurious poles. To reduce the number of spurious poles, the algorithm in [59,60] uses heuristics to determine whether the prescribed degree for $q$ was not too large and lowers it if necessary.

Let $\tilde{u} = \sum_{i=1}^m 1_{[t_{i-1}, t_i]} \tilde{u}_i$ be an approximate solution of the ODE, where $\tilde{u}_i = \tilde{a}_i^0 + 2 \sum_{k=1}^{N_i - 1} \tilde{a}_k^i T_k^i$ and $(\tilde{t}_i)_{i=0}^m$ is a partition of $[0, 1]$. The idea is to use $\tilde{u}$ to
approximate the complex singularities of the (true) solution \( u^* \), namely by constructing rational interpolants for each \( \tilde{u}_i \). Consider any \( 1 \leq i \leq m, 1 \leq j \leq n \), and initialize \( N_q = 1 \). Then follow the procedure as described below:

1. Compute an approximate rational interpolant for \( [u^*_i]_j \) of type \( \left( \left\lfloor \frac{2 \tilde{N}_i}{3} \right\rfloor, N_q \right) \) with \( K = \left\lfloor \frac{2 \tilde{N}_i}{3} \right\rfloor + N_q \) by using the approximate solution \( \tilde{u}_i \). The specific choices for the parameters are motivated in Remark 2.4.1.

2. Compute the absolute value, denoted by \( \Delta \), of the difference of the approximate size of the Bernstein ellipse of \( [u^*_i]_j \) and the decay rate of \( [\tilde{a}^i]_j \). The decay rate of \( [\tilde{a}^i]_j \) is estimated by using the least-square method to find the best line through the data points

\[
\left\{ \left( k, \log \left| [\tilde{a}^i_k]_j \right| \right) : 1 \leq k \leq \tilde{N}_i - 1, \left| [\tilde{a}^i_k]_j \right| > 10^{-16} \right\}.
\]

The decay rate is then approximated by \( e^{-s} \), where \( s \) is the slope of this line. In particular, \( \Delta = |\nu_e - e^{-s}| \), where \( \nu_e \) is defined in (2.19).

3. If \( \Delta < 0.05 \), then the approximation of the relevant singularities is deemed sufficiently accurate and we terminate the procedure. Otherwise, if \( N_q < \left\lfloor \frac{\tilde{N}_i}{3} \right\rfloor \), we increase \( N_q \) by one and return to step 1. If \( N_q = \left\lfloor \frac{\tilde{N}_i}{3} \right\rfloor \), the approximation of the singularities was unsuccessful and the program is terminated.

The significance of \( \Delta \) and the specific choice for the tolerance and stopping criteria are explained in Remark 2.4.2.

Remark 2.4.1. The value for \( K \) in step 1 is the smallest value for which \( \left\lfloor \frac{2 \tilde{N}_i}{3} \right\rfloor \) is guaranteed to admit an exact solution. The motivation for choosing \( N_p = \left\lfloor \frac{2 \tilde{N}_i}{3} \right\rfloor \) is that if one expects the existence of complex singularities (which we generally do) one should choose \( N_p < \tilde{N}_i - 1 \), since the rational interpolation scheme would yield \( p = [\tilde{u}_i]_j \) and \( q \equiv 1 \) for \( N_p \geq \tilde{N}_i - 1 \). At the same time, \( N_p \) should be chosen sufficiently large in order for the rational interpolant to be an accurate approximation of \( [u^*_i]_j \). The specific choice \( N_p = \left\lfloor \frac{2 \tilde{N}_i}{3} \right\rfloor \) is based on experimentation and the suggestions in [79].

Remark 2.4.2. The quantity \( \Delta \) defined in step 2 is used to assess the accuracy of the approximation of the relevant singularities. Indeed, if the approximation of the relevant singularities is accurate, then \( \Delta \) should be relatively small by Proposition 3.2.2. In practice, \( \Delta \) varied at best between 0.01 and 0.05 which motivated the choice for the tolerance in step 3. Furthermore, the rational interpolants were constructed by using approximate solutions \( \tilde{u} \) defined on relatively fine grids with high decay on each subdomain. Hence we expected a relatively small number of complex singularities per subdomain. This motivated the choice for the stopping criterion \( N_q = \left\lfloor \frac{\tilde{N}_i}{3} \right\rfloor \).
2.5 Applications: periodic and heteroclinic orbits in the Lorenz system

In this section we demonstrate the effectiveness of domain decomposition by using the proposed method to validate periodic and heteroclinic orbits in the Lorenz system which we were not able to validate without decomposition of the domain. In Section 2.5.2 we consider the validation of a periodic orbit on the Lorenz attractor for which the procedure in Section 2.3 is directly applicable. In Section 2.5.3 we consider a family of periodic orbits near a homoclinic orbit and in Section 2.5.4 we validate a heteroclinic orbit. In the latter two cases the procedure in Section 2.3 cannot be applied directly and needs some modifications, illustrating both the limitations and the flexibility of the method.

2.5.1 Main algorithm

First we describe the main procedure used for validating solutions of (2.1):

1. Compute an approximate zero \( \tilde{x} \) of \( F \) with respect to some grid \( (\tilde{t}_i)_{i=0}^m \) and approximate the complex singularities of \( u^* \) (the exact solution of the ODE) as described in Section 2.4.2.

2. Choose the number of domains \( m \) and use the procedure in Section 2.4.1 to determine a grid \( (t_i)_{i=0}^m \) with uniform decay on each subdomain. The number of domains \( m \) needs to be chosen in such a way that \( \max \{ t_i - t_{i-1} : 1 \leq i \leq m \} \) is sufficiently small and the number of modes \( N_i \), as determined below, is sufficiently large. In practice, we chose \( m \) by experimentation (see Section 2.5.2 for an example in which we validated a periodic orbit for different \( m \)).

3. Construct an approximate solution \( \hat{x} \) on the new grid \( (t_i)_{i=0}^m \). The number of modes \( N_i \) on each subdomain \( [t_{i-1}, t_i] \) is chosen in such a way that \( |a_k^i|_\infty < 10^{-16} \) for \( k \geq N_i \).

4. Determine weights \( (\nu_i)_{i=1}^m \) for which validation is feasible. We have chosen to fix one weight \( \nu > 1 \) and set \( \nu_i = \nu \) for all \( 1 \leq i \leq m \), since by construction the decay-rates are the same on all subdomains. Furthermore, \( \nu \) is determined by first computing an initial guess \( \nu_0 \), as explained in Remark 2.5.1 below, and checking whether validation is feasible with \( \nu = \nu_0 \). This is accomplished by computing the \( Y \) and \( Z \)-bounds as defined in Section 2.6 and constructing the radii-polyomials (without interval arithmetic). Subsequently, we try to determine an interval on which all the radii-polyomials are negative. If we do not find such an interval (i.e. validation is not feasible), then we keep decreasing \( \nu \) (as long as \( \nu > 1 \)) until validation is feasible.

5. Construct the radii-polyomials with interval arithmetic and determine an interval \( I_{m,\nu} \) at which they are all negative.
Remark 2.5.1. The initial guess $\nu_0$ in step 3 is determined by a heuristic procedure that is based on analyzing the bounds $Y_{i,j}$ as stated in Proposition 2.6.1 in Section 2.6. The idea is to choose $\nu_0 > 1$ such that

$$\frac{t_i - t_{i-1}}{2\omega} \sum_{k=N_i}^{N_g(N_i-1)+1} \left| \left[ c_{k-1}^i(\hat{a}) - c_{k+1}^i(\hat{a}) \right] \frac{\nu_0^k}{k} \right| \leq \epsilon, \quad \forall \ 1 \leq i \leq m, \ 1 \leq j \leq n,$$

where $\epsilon > 0$ is a prescribed tolerance which we set equal to $10^{-14}$ in our algorithm. A rather rough estimation yields

$$\frac{t_i - t_{i-1}}{2\hat{\omega}} \sum_{k=N_i}^{N_g(N_i-1)+1} \left| \left[ c_{k-1}^i(\hat{a}) - c_{k+1}^i(\hat{a}) \right] \frac{\nu_0^k}{k} \right| \leq \frac{h\nu_0}{\hat{\omega}N_i} \sum_{k=N_i-1}^{N_g(N_i-1)+2} \left| \left[ c_k^i(\hat{a}) \right] \frac{\nu_0^k}{k} \right|,$$

where $h := \max \{t_i - t_{i-1} : 1 \leq i \leq m\}$. Note that $[\hat{a}_k^i]_j$ and $[c_k^i(\hat{a})]_j$ are both of order $O(\nu_e^{-k})$, where $\nu_e$, see (2.19), is known by construction of the grid. Therefore, assuming that $[\hat{a}_0^i]_\infty$ is roughly of the same order on all subdomains, we anticipate that the number of modes per subdomain will be fairly uniformly distributed, say $N_i \approx \bar{N}$, where $\bar{N}$ is the (rounded) average number of modes per subdomain. Altogether, we expect the quantity

$$\frac{h\nu_e}{\hat{\omega}N} \sum_{k=N-1}^{N_g(\bar{N}-1)+2} \left( \frac{\nu_0}{\nu_e} \right)^k = \frac{h\nu_e}{\hat{\omega}N \left( 1 - \frac{\nu_0}{\nu_e} \right)} \left( \left( \frac{\nu_0}{\nu_e} \right) \bar{N} - \left( \frac{\nu_0}{\nu_e} \right)^{N_g(\bar{N}-1)+3} \right)$$

(2.23)

to provide a reasonable estimate for the order of magnitude of (2.22). Moreover, since we need to choose $\nu_0 < \nu_e$ and $\bar{N}$ is relatively small compared to $N_g(\bar{N}-1)+3$, one can approximate (2.23) by

$$\frac{h\nu_e}{\hat{\omega}N \left( 1 - \frac{\nu_0}{\nu_e} \right)} \left( \frac{\nu_0}{\nu_e} \right)^\bar{N}.$$

Hence we have chosen to determine $\nu_0$ by setting the latter quantity equal to $\epsilon$.

2.5.2 Periodic orbit on the Lorenz-attractor

We have successfully applied our method to validate a periodic orbit of period $L \approx 25.0271$ in the Lorenz system for the classical parameter values. We remark that validation was not feasible without decomposition of the domain. More precisely, the procedure described in Section 2.5.1 failed for $m = 1$, i.e., with a single domain. The main obstruction to using just one domain was the need for a large number of modes to accurately approximate the orbit, which caused the bounds related to the tail of the Chebyshev approximation to be (too) large. We should mention that it is feasible to validate this periodic orbit using a Fourier basis (and hence
Figure 2.4: The approximate complex singularities of the validated periodic orbit. The complex singularities were computed with the procedure described in Section 2.4.2. However, Fourier series can be used for problems with periodic boundary conditions only. Furthermore, the number of Fourier modes needed is comparable to the total number of modes in our domain decomposition method, and the latter is readily amenable to general (non-periodic) boundary conditions.

We have reported the computational results in Table 2.1. As expected, the size of the Bernstein-ellipses $\nu_e$ (as defined in (2.19)) increases and the (rounded) average number of modes $\bar{N}$ decreases, whenever the number of subdomains $m$ is increased. Moreover, as long as the decrease of $\bar{N}$ outweighs the increase of $m$, the dimension of $X^N_\nu$ decreases, thereby making the proof computationally more efficient. In particular, $m = 34$ was the computationally most efficient choice. We remark, however, that no attempt was made to optimize $\dim X^N_\nu$ for fixed $m$. It may be possible to validate the orbit by using a significantly smaller number of modes $N_i$ per subdomain, i.e., by relaxing the requirement that $|a_k^i|_\infty < 10^{-16}$ for $k \geq N_i$. Finally, for each $m$ the initial guess $\nu_0$ for $\nu$ was slightly too large and was lowered by 0.01 in order to make validation feasible.

The approximations of the complex singularities of the orbit are shown in Figure 2.4. Note that the relevant singularities, i.e., the ones which determine the size of the smallest Bernstein ellipse, were fairly uniformly distributed. As a consequence, the resulting grids look close to uniform at first glance, as can be seen in Figure 2.5a. However, we stress that our method for distributing the grid points based on the location of the complex singularities significantly improves the computational efficiency compared to choosing a uniform grid. To illustrate this, notice the dramatic decrease in the dimension of $X^N_\nu$ as we proceed from $m = 33$ to $m = 34$, i.e., we add one grid point. This is caused by a very subtle redistribution of the grid-points, as shown in Figure 2.5a, which resulted in a relatively large
increase in the decay-rates from 1.68 to 1.87, see Table 2.1.

Indeed, at first sight the grids appear to be very similar and it is unclear how the decay-rates could have increased that much. To get some insight, we have depicted two seemingly similar subdomains \([t_{32}, 1]\) and \([\tau_{33}, 1]\) in Figure 2.5b, where \((t_i^{33})_{i=0}^{33}\) and \((\tau_i^{34})_{i=0}^{34}\) denote the grid-points for \(m = 33\) and \(m = 34\), respectively. The grid-points \(t_{32}\) and \(\tau_{33}\) are so close to each other that the sizes of the Bernstein ellipses associated to \([t_{32}, 1]\) and \([\tau_{33}, 1]\) are determined by the same pair of complex singularities. Nevertheless, the subtle movement of \(\tau_{33}\) to the right was sufficient to cause the observed increase in the decay-rates. To see this, recall that the computation of the size of the Bernstein ellipses involves a rescaling to \([-1, 1]\), as explained in Section 2.4.1. This rescaling contributes to the increase in the decay-rates. We note that in other regions in the grid, the redistribution of the grid points (when adding a grid point) leads to a change in which pole determines the size of the Bernstein-ellipse (for some domain). The combination of delicate shifts of all the grid points together leads to the major improvement in the (uniform) decay rate.

The results highlight the effectiveness of the proposed method for performing domain decomposition: the global improvement of the decay-rates due to the subtle repositioning of the grid-points could not have been achieved by merely using uniform grids.

### 2.5.3 Family of periodic orbits near a homoclinic connection

In this section we will validate periodic orbits close to the homoclinic orbit to the origin as \(\rho \downarrow \rho_{\text{hom}} \approx 13.926557407\). The map \(F\) as defined in Definition 2.3.1, however, has to be slightly adapted to accomplish this goal. The reason \(F\) has to be adapted can be seen in Figure 2.6, which depicts the dependency of \(L\) on \(\rho\). In particular, note that the bifurcation curve is almost vertical near \(\rho_{\text{hom}}\). Therefore, \(DF_N(\hat{x})\) is close to singular near the critical parameter value \(\rho_{\text{hom}}\). Consequently, the approximate inverse \(A_N\) of \(DF_N(\hat{x})\) is badly conditioned near \(\rho_{\text{hom}}\), which

<table>
<thead>
<tr>
<th>(m)</th>
<th>(\tilde{N})</th>
<th>(\dim \mathcal{X}_N^\nu)</th>
<th>(\nu)</th>
<th>(\nu_e)</th>
<th>(I_{m,\nu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>79</td>
<td>7618</td>
<td>1.1148</td>
<td>1.6413</td>
<td>([4.5581 \cdot 10^{-10}, 1.0217 \cdot 10^{-7}])</td>
</tr>
<tr>
<td>33</td>
<td>76</td>
<td>7498</td>
<td>1.1278</td>
<td>1.6837</td>
<td>([3.1192 \cdot 10^{-10}, 1.5371 \cdot 10^{-7}])</td>
</tr>
<tr>
<td>34</td>
<td>63</td>
<td>6433</td>
<td>1.1432</td>
<td>1.8749</td>
<td>([3.8158 \cdot 10^{-10}, 1.1950 \cdot 10^{-7}])</td>
</tr>
<tr>
<td>35</td>
<td>61</td>
<td>6457</td>
<td>1.1529</td>
<td>1.9056</td>
<td>([3.3529 \cdot 10^{-10}, 1.0320 \cdot 10^{-7}])</td>
</tr>
<tr>
<td>36</td>
<td>61</td>
<td>6610</td>
<td>1.1564</td>
<td>1.9115</td>
<td>([3.1282 \cdot 10^{-10}, 1.1097 \cdot 10^{-7}])</td>
</tr>
</tbody>
</table>

Table 2.1: Numerical results for a validated periodic orbit of period \(L \approx 25.0271\) in the classical Lorenz system. In each case the number of modes \(N_i\) per domain was approximately the same. The number \(\tilde{N}\) denotes the (rounded) average number of modes per domain.
Figure 2.5: (a) A plot of the two grids \((i,t_i)_{i=0}^{33}\) and \((i,\tau_i)_{i=0}^{34}\) corresponding to \(m = 33\) and \(m = 34\), respectively. (b) A plot of the grid-points \(t_{32}, \tau_{33}\) and the complex singularities (colored in red) which determine the size of the Bernstein ellipses associated to \([t_{32}, 1]\) and \([\tau_{33}, 1]\).  

causes the estimates in Proposition 2.6.8 to blow up, which in turn will obstruct the validation process.

The latter problem can be solved by adding an additional equation to \(F\) and including the parameter \(\rho\) as an additional variable to solve for. We adapt the method in Section 2.3 as follows:

- Include \(\rho\) as an additional variable in \(\mathcal{X}_\nu\), i.e., set \(\mathcal{X}_\nu := \mathbb{R} \times \mathbb{R} \times \prod_{i=1}^{m} \ell^{1}_{(\nu,n)}\) and write \(x = (\rho, \omega, a^1, \ldots, a^m) \in \mathcal{X}_\nu\).

- Define the norm and projections on \(\mathcal{X}_\nu\) in the same way as before by including an additional projection \(\Pi^{-1}_{-1}: \mathcal{X}_\nu \rightarrow \mathbb{R}\) onto the parameter space defined by \(\Pi_{-1}(\rho, \omega, a^1, \ldots, a^m) := \rho\).

- Define \(\tilde{F}: \mathcal{X}_\nu \rightarrow \mathcal{X}_\nu\) by
  \[
  \tilde{F}(x) := (f_{-1}(\rho, \omega, a^1, \ldots, a^m), f_0(a^1), f_1(\rho, \omega, a^1, a^m), f_2(\rho, \omega, a^1, a^2), \ldots, f_m(\rho, \omega, a^{m-1}, a^m)),
  \]
  where \(f_0, \ldots, f_m\) are defined as before, and we choose
  \[
  f_{-1} := \langle V_0, \Pi_N(x) - U_0 \rangle,
  \]
  where \(V_0\) is approximately tangent to the solution curve \((\rho, \phi(\rho))\) of \(F_N\), and \(U_0\) is a “predictor” for the next point on the solution curve.

The corresponding modifications to the bounds \(Y\) and \(Z\) are described in Section 2.6.3.
Results

To examine the performance of the proposed method we first determined how far we could push the period by using only one domain. Next, we extended the result by using domain decomposition. In particular, we validated a long periodic orbit of period $L \approx 100.2254$ which revealed a limitation of the proposed method. In fact, there the standard algorithm breaks down in two spots.

First, it was not feasible to determine a grid by using the procedure in Section 2.4, since in the region where the orbit is flat (i.e. near the equilibrium at the origin in phase space, see Figure 2.3) we were not able to compute accurate approximations of the complex singularities. A likely reason for this is that the complex singularities in this region are located too far away from the real axis (i.e. there are no “nearby” poles).

Second, after fixing the grid in the flat part of the solution in an ad-hoc manner (discussed below), the number of modes $N_i$ in this part of the grid, as determined via the procedure in Section 2.5.1 was very small. To see why the use of such a small number of modes is an obstruction, recall that the approximate inverse $A$, as defined in Definition 2.3.6, was constructed under the assumption that $\omega k a_i^k$ is the dominant term in $(f_i (\omega, a^{-1}, a^i))_k$ for $k \geq N_i$ in a small neighborhood of the numerical approximation. The latter assumption, however, is only satisfied if the truncation dimensions $N_i$ are sufficiently large and $\max \{t_i - t_{i-1} : 1 \leq i \leq m\}$ is sufficiently small (see Definition 2.3.1). Consequently, in order to validate the flat part of the orbit (where a small number of modes per subdomain is used), one
Table 2.2: Numerical results for two periodic orbits near the homoclinic connection. The periodic orbit of period $L \approx 100.2554$ was in both cases validated on a grid for which (the same) six subdomains were used to approximate the non-flat part of the orbit.

<table>
<thead>
<tr>
<th>$L$</th>
<th>$m$</th>
<th>$\dim \mathcal{A}^N_{\nu}$</th>
<th>$I_{m,\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5473</td>
<td>1</td>
<td>566</td>
<td>$[4.4568 \cdot 10^{-11}, 8.4403 \cdot 10^{-6}]$</td>
</tr>
<tr>
<td>100.2554</td>
<td>7</td>
<td>5894</td>
<td>$[1.5186 \cdot 10^{-11}, 7.4915 \cdot 10^{-8}]$</td>
</tr>
<tr>
<td>100.2554</td>
<td>506</td>
<td>8441</td>
<td>$[1.5174 \cdot 10^{-11}, 4.2914 \cdot 10^{-6}]$</td>
</tr>
</tbody>
</table>

needs to ensure that the grid is sufficiently fine there.

To validate the long periodic orbit we constructed a grid which was uniform in the region where the orbit is flat, and outside this region (where the number of modes per subdomain were relatively large) the grid-points were placed by using the complex singularities as described in Section 2.4. We remark that another strategy for resolving the above issue is to use only one domain for the flat part of the orbit, and to artificially increase the number of modes on this subdomain by padding with zeros. We have succeeded in validating the orbit in this way as well. The results are reported in Table 2.2. In particular, in the case $m = 7$ we used one subdomain with 1800 modes (of which only the first 136 were nonzero) to approximate the flat part of the orbit. In the case $m = 506$ we used 500 equally spaced subdomains each using (on average) about five modes to represent the flat part of the orbit.

By adapting the algorithm, we are thus able to validate very long orbits near the homoclinic connection. We conclude this section by remarking on two possible improvements to the domain decomposition technique.

**Remark 2.5.2.** The results show that the proposed method is not directly applicable for validating orbits which exhibit slow-fast behavior on different time-scales. In this particular case, a more effective approach for validating the long periodic orbit would be to avoid “numerical integration” of the slow passage and to analyze the (relatively simple) dynamics near the equilibrium by other means (normal forms, lambda lemma, etc.).

**Remark 2.5.3.** For this particular problem, distribution of the grid-points based on the location of the complex singularities is not an efficient choice, since our domain decomposition algorithm will concentrate most of the grid-points outside the region where the orbit is flat. Indeed, (in general) our domain decomposition algorithm will yield relatively large subdomains in regions where the complex singularities are located far away from the real-axis. This can obstruct the validation process as $\max \{t_i - t_{i-1} : 1 \leq i \leq m\}$ might be too large. To resolve this issue, one can try to improve the domain decomposition algorithm by incorporating constraints on the maximal distance between successive grid-points.
2.5.4 Heteroclinic orbit

To show that our method is applicable to more general BVPs than just periodic boundary conditions, we consider the validation of a transverse heteroclinic orbit from \( q^+ = (\sqrt{\beta}(\rho - 1), \sqrt{\beta}(\rho - 1), \rho - 1) \) to the origin for the classical parameter values in the Lorenz system. Both the origin and \( q^+ \) are hyperbolic, and \( \dim(W^s(0)) = \dim(W^u(q^+)) = 2 \). In particular, the transversality condition \( n_u + n_s = n + 1 \), where \( n_u = \dim(W^u(q^+)) \), \( n_s = \dim(W^s(0)) \), is satisfied.

The idea is to set up a suitable BVP which characterizes the heteroclinic orbit, and to adjust the method in Section 2.3 accordingly in order to solve the BVP. A heteroclinic orbit from \( q^+ \) to the origin is characterized by

\[
\begin{cases} 
\frac{du}{dt} = Lg(u), & t \in [0, 1], \\
u(0) = P(\alpha), & \alpha \in V_u, \\
u(1) = Q(\phi), & \phi \in V_s,
\end{cases}
\]

where \( L > 0 \) is a fixed integration time and \( P : V_u \subset \mathbb{R}^2 \rightarrow W^u_{loc}(q^+) \), \( Q : V_s \subset \mathbb{R}^2 \rightarrow W^s_{loc}(0) \) are local parameterizations of \( W^u_{loc}(q^+) \) and \( W^s_{loc}(0) \), respectively. We have used the parameterization method developed in \([18, 49, 74]\) to explicitly compute \( P \) and \( Q \).

The idea of the computational method developed in \([18, 49, 74]\) is to construct \( P \) by expanding it as a power series, and by requiring that it conjugates the unstable part of the linearized dynamics around the origin with the dynamics on \( W^u_{loc}(q^+) \). The parameterization \( Q \) is obtained similarly. The method yields approximate parameterizations \( P_{N_u} \) and \( Q_{N_s} \), where \( N_u, N_s \in \mathbb{N} \) are the degrees up to which the power series are computed, and establishes the existence of exact parameterizations \( P \) and \( Q \) via a rigorous numerical scheme. In particular, the procedure provides rigorous error bounds \( \delta_u, \delta_s > 0 \) such that

\[
\|P - P_{N_u}\|_\infty \leq \delta_u, \quad \|Q - Q_{N_s}\|_\infty \leq \delta_s.
\]

Since heteroclinic orbits are invariant under translations in time, we need to introduce a phase condition to remove this extra degree of freedom. This can be accomplished by, roughly speaking, restricting \( P \) or \( Q \) to a domain of one dimension less. We have used the same phase condition as in \([49]\): let \( \Theta_\mu : S^1 \rightarrow V_s \) be the embedding of the unit circle into \( V_s \) defined by

\[
\Theta_\mu(\phi) := \mu(\cos \phi, \sin \phi),
\]

where \( \mu > 0 \) is sufficiently small, and consider the following equivalent formulation of (2.24):

\[
\begin{cases} 
\frac{du}{dt} = Lg(u), & t \in [0, 1], \\
u(0) = P(\alpha), & \alpha \in V_u, \\
u(1) = Q \circ \Theta_\mu(\phi), & \phi \in V_s.
\end{cases}
\]
The procedure in Section 2.3, however, needs to be modified before it can be applied to (2.25). First, note that we fix the integration time $L$. Furthermore, the parameterization variables $\phi$ and $\alpha$ have to be treated as unknown variables. Therefore, in order to solve (2.25) we modify the procedure in Section 2.3 as follows:

- Set $X_\nu := S^1 \times V_\nu \times \prod_{i=1}^m \ell_{(v_i,n)}$ and write $x = (\phi, \alpha, a^1, \ldots, a^m)$.
- Adapt the set-up described in Section 2.3.2 by replacing $\Pi_0$ with projections $\Pi_{0,j} : X_\nu \rightarrow \mathbb{R}$, where $1 \leq j \leq 3$, defined by
  \[ \Pi_{0,1}(\phi, \alpha, a^1, \ldots, a^m) := \phi, \quad \Pi_{0,j}(\phi, \alpha, a^1, \ldots, a^m) = [\alpha]_{j-1}, \]
  for $j \in \{2, 3\}$.
- Define $F : X_\nu \rightarrow \mathbb{R}^3 \times \prod_{i=1}^m \ell_{(\nu_i,n)}$ analogously as in Definition 2.3.1 by incorporating the modified boundary conditions into $f_0$ and $(f_1)_0$:
  \[ F(x) := (f_0(\phi, a^m), f_1(\alpha, a^1), f_2(a^1, a^2), \ldots, f_m(a^{m-1}, a^m)), \]
  where
  \[ f_0(\phi, a^m) := a_m^m + 2 \sum_{k=1}^{\infty} a_k^m - Q \circ \Theta_\mu(\phi), \]
  \[ (\Pi_1 F(x))_0 = a_1^1 + 2 \sum_{k=1}^{\infty} (-1)^k a_k^1 - P(\alpha), \]
  \[ (\Pi_i F(x))_k = ka_k^i - \frac{L(t_i - t_{i-1})}{4} (c_{k-1}^i - c_{k+1}^i), \quad k \in \mathbb{N}, 1 \leq i \leq m. \]
- Define the finite dimensional reduction $F_N : \mathbb{R}^{n(1+\sum_{i=1}^m N_i)} \rightarrow \mathbb{R}^{n(1+\sum_{i=1}^m N_i)}$ of $F$ by $F_N(x) := \Pi_N F(x)$, and by replacing $P, Q$ with $P_{N_u}, Q_{N_u}$, respectively.
- Define $A$ and $\hat{A}$ as before without the factors $\frac{1}{\omega}$ and $\hat{\omega}$, respectively.

The corresponding modifications to the estimates $Y$ and $Z$ are described in Section 2.6.3

**Results**

We have successfully validated a connecting orbit from $q^+$ to the origin by using the procedure described in Section 2.5.1. The integration time was $L = 30$. The parameters used for approximating the stable and unstable manifolds were $N_u = 15$, $N_s = 25$, $\mu = 0.4$, and $r^* = 10^{-6}$. The meaning of $r^*$ is explained in Section 2.6.3. The corresponding error-bounds for the parameterizations were $\delta_u \leq 4.6847$. 

52
Table 2.3: Numerical results for the connecting orbit from $q^+$ to the origin. The interval $I_{m,\nu}$ is the set of admissible radii on which the radii-polynomials were proven to be strictly negative.

![Table 2.3: Numerical results for the connecting orbit from $q^+$ to the origin. The interval $I_{m,\nu}$ is the set of admissible radii on which the radii-polynomials were proven to be strictly negative.](image)

10^{-12} \text{ and } \delta_s \leq 5.9717 \cdot 10^{-15}. \text{ We have kept the size of } W^u_{loc}(q^+) \text{ small so that the orbit was relatively “long” and sufficiently complicated to test the domain decomposition method. The computational results are reported in Table 2.3.}

The complex singularities and the corresponding grid are shown in Figures 2.7a and 2.7b respectively. Figure 2.7a shows that the complex singularities move closer to the real axis as the orbit spirals away from $q^+$ up until the point at which the orbit travels to the origin in (roughly) a straight line in phase space (see Figure 2.1). In this last part of the orbit there appear to be no complex singularities close to the real-axis. These observations are reflected in the distribution of the grid-points as shown in Figure 2.7b: the distance between successive grid-points decreases as the orbit spirals away from $q^+$, except for the distance between the second to last grid point and the last one, which is substantially larger.

2.6 The estimates needed to prove contraction

In this section we give explicit expressions for the bounds $Y$ and $Z$ in Theorem 2.3.10. We focus primarily on periodic boundary conditions. Additionally, we indicate where (and which) changes are in order for more general types of boundary
conditions. Explicit examples of such generalizations are discussed in Section 2.6.3 which deals with the modifications of the estimates that arise in the applications in Sections 2.5.3 and 2.5.4.

2.6.1 Computation of the $Y$-bounds

Proposition 2.6.1. The bounds

$$Y_0 := |\Pi_0 A_N F_N (\hat{x})|,$$

$$Y_{i,j} := \frac{t_i - t_{i-1}}{2\omega} \sum_{k=N_i}^{N_g(N_i-1)+1} \left| c_{k-1}(\hat{\alpha}) - c_{k+1}(\hat{\alpha}) \right| \nu_i^{k} + \|\Pi_{i,j} A_N F_N (\hat{x})\|_{\nu_i},$$

where $1 \leq i \leq m$, $1 \leq j \leq n$, satisfy (2.12).

Proof. Let $1 \leq i \leq m$, $1 \leq j \leq n$ be arbitrary and note that

$$|\Pi_0 (T (\hat{x}) - \hat{x})| \leq |\Pi_0 A (F (\hat{x}) - F_N (\hat{x}))| + |\Pi_0 AF_N (\hat{x})|,$$

$$\|\Pi_{i,j} (T (\hat{x}) - \hat{x})\|_{\nu_i} \leq \|\Pi_{i,j} A (F (\hat{x}) - F_N (\hat{x}))\|_{\nu_i} + \|\Pi_{i,j} AF_N (\hat{x})\|_{\nu_i}.$$  \hfill (2.26)

Next, observe that the only nonzero components of $F(\hat{x}) - F_N(\hat{x})$ are

$$(\Pi_{i,j} (F(\hat{x}) - F_N(\hat{x})))_k = -\frac{(t_i - t_{i-1})}{4} \left[ c_{k-1}(\hat{\alpha}) - c_{k+1}(\hat{\alpha}) \right]_j$$

for $N_i \leq k \leq N_g(N_i-1)+1$. Hence the result follows from (2.26) and (2.27). \qed

2.6.2 Computation of the $Z$-bounds

Let $x_1, x_2 \in B_r(0)$, $r > 0$ be arbitrary and recall the factorization in (2.11). We shall compute bounds $Z_{i,j}(r)$ and $Z_{0}(r)$ satisfying (2.13) and (2.15), respectively, by estimating the two terms in (2.11) separately. Throughout this section we write $x_1 = rv$ and $x_2 = rw$, where $v = (\omega, v^1, \ldots, v^m)$, $w = (\omega, w^1, \ldots, w^m) \in B_1(0)$.

We start by computing a bound for $(I - A\hat{A}) x_2$. To accomplish this we first state a result about the norm of linear operators $C$ on $\mathcal{X}_r^N$, for which we introduce the notation

$$\Pi_0 C(\omega, 0) = C_0^0 \omega \quad \Pi_0 C(0, a) = \sum_{i,j,k} (C^0_a)_{ijk} a_{ijk}$$

$$(\Pi_a C(\omega, 0))_{ij,k} = (C^{0}_a)_{ij,k} \omega \quad \Pi_a C(0, a)_{ijk} = \sum_{i,j,k} (C^a_a)_{ijk} a_{ijk}$$

where $i, \tilde{i} = 1, \ldots, m$ and $j, \tilde{j} = 1, \ldots, n$ and $k, \tilde{k} = 0, \ldots, N_i - 1$ refer to the notation $a_{ijk} = ([a_{k}^l]_j$ for the Chebyshev coefficients introduced in Section 2.3.1.
Lemma 2.6.2. Suppose $C : X^N_{\nu} \to X^N_{\nu}$ is a linear operator. Using the above notation, we define

$$\eta_{ij} := \| (C_{a}^0)_{ij} \|_{\nu, i}$$
$$\mu_{ij}^{\tilde{i} \tilde{j}} := \| (C_{a}^0)^{\tilde{i} \tilde{j}} \|_{\nu, i}$$
$$\xi^{\tilde{i} \tilde{j} \tilde{k}}_{ij} := \| (C_{a}^0)^{\tilde{i} \tilde{j} \tilde{k}} \|_{\nu, i}$$
$$\xi_{ij}^{\tilde{i} \tilde{j}} := \| \xi^{\tilde{i} \tilde{j}} \|_{\nu, i}.$$ 

Then

$$\| \Pi_0 C \|_{\mathcal{B}(X^N_{\nu}, \mathbb{R})} \leq |C^0_0| + \sum_{i=1}^{m} \sum_{j=1}^{n} \mu_{ij}^{\tilde{i} \tilde{j}},$$
$$\| \Pi_{i,j} C \|_{\mathcal{B}(X^N_{\nu}, \ell^1_{\nu, i})} \leq \eta_{ij} + \sum_{i=1}^{m} \sum_{j=1}^{n} \xi^{\tilde{i} \tilde{j}}_{ij}.$$ 

Proof. This follows from writing out the definitions of the norms.

Remark 2.6.3. Explicit expressions for $\mu_{ij}^{\tilde{i} \tilde{j}}$ and $\xi^{\tilde{i} \tilde{j} \tilde{k}}_{ij}$ can be obtained by using Lemma 2.2.5.

We can now compute a bound for $(I - A\hat{A}) x_2$.

Lemma 2.6.4. Let $1 \leq i \leq m$, $1 \leq j \leq n$, and let $h_0$ and $h_{i,j} > 0$ denote the bounds

$$\| \Pi_0 (I_N - A_N DF_N (\hat{x})) \|_{\mathcal{B}(X^N_{\nu}, \mathbb{R})} \leq h_0,$$
$$\| \Pi_{i,j} (I_N - A_N DF_N (\hat{x})) \|_{\mathcal{B}(X^N_{\nu}, \ell^1_{\nu, i})} \leq h_{i,j},$$

provided by Lemma 2.6.2, where $I_N$ is the identity on $X^N_{\nu}$. Then

$$\| \Pi_0 (I - A\hat{A}) x_2 \|_{\nu, i} \leq h_{0} r, \quad \| \Pi_{i,j} (I - A\hat{A}) x_2 \|_{\nu, i} \leq h_{i,j} r,$$

for all $1 \leq i \leq m$, $1 \leq j \leq n$.

Proof. It suffices to observe that

$$I - A\hat{A} = \Pi_N \left( I - A\hat{A} \right) = I_N - A_N DF_N (\hat{x}),$$

since the tails of $A$ and $\hat{A}$ are exact inverses of each other.

The analysis of the second term in (2.11) is more complicated and requires one to analyze the infinite dimensional map $F$ in more detail. Note that

$$\left( DF (\hat{x} + rv) - \hat{A} \right) w = \left. \frac{d}{d\tau} \right|_{\tau=0} \left( F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w) \right) - \hat{A}_\infty w,$$

(2.28)
where $\hat{A}_\infty = (I - \Pi_N) \hat{A}$, since $\Pi_N \hat{A} = DF_N (\hat{x})$. Furthermore, a straightforward computation shows that

$$
\left( \frac{d}{d\tau} \right)_{\tau=0} \Pi_i (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w))_0 = 2 \left( \sum_{k=N_i}^{\infty} (-1)^k w^i_k - \sum_{k=N_i-1}^{\infty} w^i_k \right),
$$

(2.29)

for $1 < i \leq m$, and

$$
\left( \frac{d}{d\tau} \right)_{\tau=0} \Pi_i (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w))_k = k \left( \omega_w v^i_k + \omega_v w^i_k \right) r
$$

$$
- \frac{t_i - t_{i-1}}{4} \left( c^i_{k-1} (\hat{a}^i + rv^i + \tau w^i) - c^i_{k-1} (\hat{a}^i + \tau \Pi_i \nu^i \nu^i) \right)
$$

$$
- c^i_{k+1} (\hat{a}^i + rv^i + \tau w^i) + c^i_{k+1} (\hat{a}^i + \tau \Pi_i \nu^i \nu^i) \right),
$$

(2.30)

for $1 \leq i \leq m$, $1 \leq k \leq N_i - 1$, while

$$
\left( \frac{d}{d\tau} \right)_{\tau=0} \Pi_i (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w) - \Pi_i \hat{A}_\infty w)_k
$$

$$
= - \frac{t_i - t_{i-1}}{4} \left( c^i_{k-1} (\hat{a}^i + rv^i + \tau w^i) - c^i_{k-1} (\hat{a}^i + rv^i + \tau w^i) \right)
$$

$$
+ k \left( \omega_w v^i_k + \omega_v w^i_k \right) r
$$

(2.31)

for $1 \leq i \leq m$, $k \geq N_i$.

We start by computing a bound for (2.29):

**Lemma 2.6.5.** Let $1 < i \leq m$, $1 \leq j \leq m$, then

$$
\left| \left( \frac{d}{d\tau} \right)_{\tau=0} \Pi_{i,j} (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w)) \right|_0 \leq \nu_i^{-N_i} + \nu_i^{-N_i-1}.
$$

**Proof.** Define $\psi_i : \ell^1_{\nu_i} \to \mathbb{R}$ by $\psi_i(x) := 2 \sum_{k=N_i}^{\infty} x_k$, and note that $\psi_i \in (\ell^1_{\nu_i})^*$ for any $1 \leq i \leq m$, since $\nu_i > 1$. Furthermore, $\|\psi_i\|_{\nu_i} = \nu_i^{-N_i}$ by Lemma 2.2.5. Therefore, one can bound the components of (2.29) by $\nu_i^{-N_i} + \nu_i^{-N_i-1}$. □

Next, we compute component-wise bounds for the convolution terms in (2.30) for an arbitrary subdomain. Since the construction of these bounds is the same for each subdomain, we will fix and omit the superscript $i$ whenever possible.
Furthermore, to avoid additional clutter we shall denote the \( j \)-th component of a sequence \( a \) by \( a_j \) instead of \( [a]_j \) whenever there is no chance of confusion.

Recall that the convolution terms are defined by
\[
c(a) = \sum_{|\alpha| \leq N_g} g_\alpha a^\alpha,
\]
where \( a^\alpha = a_1^{\alpha_1} \cdots a_n^{\alpha_n} \), \( a \in \ell^1_{(\nu_i, n)} \) and \( N_g \in \mathbb{N} \) is the degree of the (polynomial) vector field. As mentioned before in the introduction, for the sake of simplicity, we shall restrict our attention to the case in which \( N_g = 2 \). In particular, note that
\[
\sum_{|\alpha| = 2} g_\alpha a^\alpha = \sum_{1 \leq l \leq s \leq n} g_{ls} a_l * a_s,
\]
where \( g_{ls} = g_{e_l + e_s} \) and \( (e_j)^n_{j=1} \) are the unit vectors in \( \mathbb{R}^n \).

A straightforward computation shows that
\[
\frac{d}{d\tau} \left|_{\tau=0} \right. \left[ c(\hat{a} + rv + \tau w) - c(\hat{a} + \tau \Pi_{N_i}(w)) \right] = \sum_{j=1}^n g_{e_j} [\tilde{w}]_j + \sum_{1 \leq l \leq s \leq n} g_{ls} \left( \tilde{w}_l * \hat{a}_s + \tilde{w}_s * \hat{a}_l + r (w_l * v_s + w_s * v_l) \right), \tag{2.32}
\]
where
\[
\tilde{w} = \begin{cases} 0_n, & 0 \leq k \leq N_i - 1, \\ w_k, & k \geq N_i. \end{cases}
\]
The following lemma is key in computing bounds for the linear terms in (2.32):

**Lemma 2.6.6.** Let \( a \in \ell^1_{\nu_i} \) be such that \( a_k = 0 \) for \( k \geq N_i \). Define \( \Psi_{a,k} : \ell^1_{\nu_i} \to \mathbb{R} \) by
\[
\Psi_{a,k}(x) := (\tilde{x} * a)_{k-1} - (\tilde{x} * a)_{k+1},
\]
for \( 1 \leq k \leq N_i - 1 \), where \( \tilde{x} \) is defined by
\[
\tilde{x} = \begin{cases} 0, & 0 \leq k \leq N_i - 1, \\ x_k, & k \geq N_i. \end{cases}
\]
Then \( \Psi_{a,k} \in (\ell^1_{\nu_i})^* \) and
\[
\|\Psi_{a,k}\|_{\nu_i}^* = \frac{1}{2} \max \left( \left\{ \nu_{i}^{-l} |a_{|k-1-l|} - a_{|k+1-l|}| \right\}_{l=N_i}^{l=k+N_i-2}, \nu_{i}^{-(k+N_i)} |a_{|N_i-2|}|, \nu_{i}^{-(k+N_i)} |a_{|N_i-1|}| \right)
\]
for \( 1 \leq k \leq N_i - 1 \).
Proof. Let \( x \in \ell^1_{\nu_i}, \ k \in \mathbb{N}_0 \) be arbitrary and observe that

\[
(\tilde{x} * a)_k = \sum_{k_1 = -N_i}^{k + N_i - 1} x_{k_1} a_{|k - k_1|},
\]

(2.33)
since \( \tilde{x}_{k_1} = 0, \ a_{k_1} = 0, \) for \( 0 \leq k_1 \leq N_i - 1 \) and \( k_2 \geq N_i, \) respectively. Next, note that \( \Psi_{a,k} \in (\ell^1_{\nu_i}) \) by Proposition 2.2.4 and

\[
\Psi_{a,k}(\varepsilon l) = \begin{cases} 
\frac{\nu_i^{-l}}{2} (a_{|k-l|} - a_{|k+1-l|}), & N_i \leq l \leq k + N_i - 2, \\
-\frac{\nu_i^{-l}}{2} a_{|k+1-l|}, & l = k + N_i - 1, \ k + N_i, \\
0, & \text{otherwise}.
\end{cases}
\]

Now use Lemma 2.2.5 to obtain the stated formula for \( \|\Psi_{a,k}\|_{\nu_i}^* \).

Corollary 2.6.7. Let \( 1 \leq i \leq m \) and \( 1 \leq k \leq N_i - 1, \) then

\[
\left| \left( \sum_{j=1}^{n} g_{e_j} [\bar{w}]_j + \sum_{1 \leq l \leq s \leq n} g_{ls} (\bar{w}_l * \hat{a}_s + \bar{w}_s * \hat{a}_l) \right)_{k-1} - \left( \sum_{j=1}^{n} g_{e_j} [\bar{w}]_j + \sum_{1 \leq l \leq s \leq n} g_{ls} (\bar{w}_l * \hat{a}_s + \bar{w}_s * \hat{a}_l) \right)_{k+1} \right|
\]

is bounded by (using the Kronecker \( \delta \))

\[
B^i_k := \delta_k, N_i - 1 \sum_{j=1}^{N_i} |g_{e_j}| + \sum_{1 \leq l \leq s \leq n} |g_{ls}| \left( \|\Psi_{\hat{a}_i,k}^*\|_{\nu_i}^* + \|\Psi_{\hat{a}_s,k}^*\|_{\nu_i}^* \right).
\]

(2.34)

Proof. This is a straightforward consequence of Lemma 2.6.6.

We are now ready to construct bounds for

\[
\|\Pi_{i,j} \Pi_N A \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2\|_{\nu_i}.
\]

(2.35)

There are two boundary conditions that we need to deal with separately, namely the phase condition and the periodicity condition (the “internal” boundary conditions between successive domains will be dealt with uniformly). We deal with these two bounds in such a way that the method can be easily adapted to deal with other boundary conditions. Hence, for the moment, assume that there exist bounds \( \Lambda_{0,1}, \ \Lambda_{0,2} > 0, \) and \( \Lambda_{1,1}, \ \Lambda_{1,2} \in \mathbb{R}_{\geq 0}^n, \) such that

\[
\left| \frac{d}{d\tau} \right|_{\tau=0} \Pi_0 (F(\hat{x} + r v + \tau w) - F_N (\hat{x} + \tau w)) \leq \Lambda_{0,1} + r \Lambda_{0,2},
\]

(2.36)

\[
\left| \frac{d}{d\tau} \right|_{\tau=0} (\Pi_1 (F(\hat{x} + r v + \tau w) - F_N (\hat{x} + \tau w)))_0 \leq \Lambda_{1,1} + r \Lambda_{1,2},
\]

(2.37)

58
for any \(v, w \in B_1(0)\). Explicit expressions for these bounds are given explicitly in Remark 2.6.9 for periodic boundary conditions.

We define \(\tilde{Z}_1 \in \mathbb{R}^{1+n} \sum_{i=1}^n N_i \) by \(\Pi_0 \tilde{Z}_1 := \Lambda_{0,1} \) and

\[
\Pi_{N_i} \tilde{Z}_1 := \left[ \frac{\Lambda_{1,1}}{4} t_i - t_0 \right]_{k=1}^{B_k} N_i , \quad \Pi_{N_i} \tilde{Z}_1 := \left[ \left( \nu_{N_i} - \nu_{N_i-1} \right) \cdot \mathbf{1}_n \right]_{k=1}^{B_k} N_i ,
\]

with \(B_k\) defined in (2.34), and we set \(Z_1 = |A_N| \tilde{Z}_1\).

**Proposition 2.6.8.** Let \(1 \leq i \leq m, 1 \leq j \leq n\), then

\[
\Pi_0 A \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2 \leq \Pi_0 Z_1 r + \gamma \Pi_0 A_N \|B(\chi_{N_i}, r)\| r^2 ,
\]

\[
\left\| \Pi_{i,j} \Pi_N A \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2 \right\|_{\nu_i} \leq \left\| \Pi_{i,j} Z_1 \right\|_{\nu_i} r + \gamma \left\| \Pi_{i,j} A_N \right\|_{B(\chi_{N_i}, r)} r^2 ,
\]

where the operator norms can be evaluated using Lemma 2.6.2 and

\[
\gamma := \max \left\{ \{\Lambda_{0,2}\} \cup \left\{ \Lambda_{1,2} + 2 (N_1 - 1) + \frac{(t_1 - t_0) \left( 2 \nu_1^2 + 1 \right)}{2 \nu_1} \sum_{|\alpha| = 2} |g_\alpha| \right\}_\infty \right\} \cup \left\{ 2 (N_i - 1) + \frac{(t_i - t_{i-1}) \left( 2 \nu_i^2 + 1 \right)}{2 \nu_i} \sum_{|\alpha| = 2} |g_\alpha| : 2 \leq i \leq m \right\}.
\]

**Proof.** First observe that

\[
\left\| \Pi_0 \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2 \right\| \leq \Lambda_{0,1} r + \Lambda_{0,2} r^2
\]

by (2.36) and

\[
\left\| \Pi_{N_i} \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2 \right\| \leq \Pi_{N_i} \left( \tilde{Z}_1 \right) r + \tilde{Z}_{2,1} r^2 , \quad (2.38)
\]

where

\[
(\tilde{Z}_{2,1})_k := \begin{cases} 
\Lambda_{1,2}, & k = 0, \\
\sum_{1 \leq l \leq s \leq n} |g_{ls}| \left( |(w_l * v_s)_{k-1}| + |(w_s * v_l)_{k-1}| + |(w_s * v_l)_{k+1}| + |(w_s * v_l)_{k+1}| \right), & 1 \leq k \leq N_i - 1,
\end{cases}
\]
Furthermore, the quadratic part of (2.38) is estimated by
\[
\|[\tilde{Z}_{2,1}]\|_{(\nu_i,n)} \leq \max_{1 \leq j \leq n} \left( [\Lambda_{1,2}]_j + 2 (N_1 - 1) + \frac{(t_1 - t_0)(2\nu_1^2 + 1)}{2\nu_1} \sum_{|\alpha|=2} \left| [g_\alpha]_j \right| \right).
\]

Similarly,
\[
\left| \Pi_{N_i} \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2 \right| \leq \Pi_{N_i} \left( \tilde{Z}_1 \right) r + \tilde{Z}_{2,i} r^2,
\]
where
\[
(\tilde{Z}_{2,i})_k :=
\begin{cases}
0_n, & k = 0, \\
k (|v_k| + |w_k|) + \frac{t_i - t_{i-1}}{4} \sum_{1 \leq l \leq s \leq n} |g_l| \left( |(w_l * v_s)_{k-1}| + |(w_s * v_l)_{k-1}| + |(w_s * v_l)_{k+1}| + |(w_l * v_s)_{k+1}| \right), & 1 \leq k \leq N_i - 1,
\end{cases}
\]
and the quadratic part is estimated by
\[
\|[\tilde{Z}_{2,i}]\|_{(\nu_i,n)} \leq \max_{1 \leq j \leq n} \left( 2 (N_i - 1) + \frac{(t_i - t_{i-1})(2\nu_i^2 + 1)}{2\nu_i} \sum_{|\alpha|=2} \left| [g_\alpha]_j \right| \right).
\]

Combining these estimates, we can now bound \( A \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2 \) as asserted.

\[\blacksquare\]

**Remark 2.6.9.** In the current setting for periodic orbits we have that
\[
\left. \frac{d}{dt} \right|_{\tau=0} \Pi_0 (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w)) = 0,
\]
\[
\left( \left. \frac{d}{dt} \right|_{\tau=0} \Pi_1 (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w)) \right)_0 = 2 \left( \sum_{k=N_1}^{\infty} (-1)^k w_k^1 - \sum_{k=N_m}^{\infty} w_k^m \right).
\]

Therefore, by the same computation as in Lemma 2.6.3, it suffices to set
\[
\Lambda_{1,1} = \left( \nu_1^{-N_1} + \nu_m^{-N_m} \right) \cdot 1_n, \quad \Lambda_{1,2} = 0_n, \quad \Lambda_{0,1} = \Lambda_{0,2} = 0.
\]

60
It remains to bound the $\ell^1_{\nu_i}$-norm of $2.31$ for $k \geq N_i$. Observe that

$$\frac{d}{d\tau} \bigg|_{\tau=0} c(\hat{a} + rv + \tau w)$$

$$= \sum_{j=1}^{n} g_j [w]_j + \sum_{1 \leq l \leq s \leq n} g_{ls} (w_l \ast \hat{a}_s + w_s \ast \hat{a}_l + r (w_l \ast v_s + w_s \ast v_l)). \quad (2.39)$$

**Lemma 2.6.10.** Let $a \in \ell^1_{\nu_i}$ be such that $a_k = 0$ for $k \geq N_i$. Define $\varphi^-_a, \varphi^+_a : \ell^1_{\nu_i} \to \mathbb{R}$ by

$$\varphi^-_a (x) := \sum_{k=N_i-1}^{\infty} (x \ast |a|)_k \frac{\nu^k_i}{k+1}, \quad \varphi^+_a (x) := \sum_{k=N_i+1}^{\infty} (x \ast |a|)_k \frac{\nu^k_i}{k-1}.$$  

Then $\varphi^-_a, \varphi^+_a \in (\ell^1_{\nu_i})^*$, and $\|\varphi^-_a\|_{\nu_i} = \frac{1}{2} \Gamma^-_a$, $\|\varphi^+_a\|_{\nu_i} = \frac{1}{2} \Gamma^+_a$, where

$$\Gamma^-_a := \max \left\{ \frac{1}{2} \sum_{k=N_i-1-l}^{N_i-1} |a|_k \frac{\nu^k_i}{k+l+1} : 1 \leq l \leq 2 (N_i - 1) \right\}$$

$$\cup \left\{ 2 |a_{N_i-1}| \frac{\nu^{N_i-1}_i}{N_i} \right\},$$

$$\Gamma^+_a := \max \left\{ \frac{1}{2} \sum_{k=N_i+1-l}^{N_i-1} |a|_k \frac{\nu^k_i}{k+l-1} : 2 \leq l \leq 2N_i \right\}.$$  

**Proof.** It follows directly from Proposition 2.2.4 that $\varphi^-_a, \varphi^+_a \in (\ell^1_{\nu_i})^*$. Next, we consider the computation of $\|\varphi^-_a\|_{\nu_i}$ (the computation of $\|\varphi^+_a\|_{\nu_i}$ is similar). Let $k \geq N_i - 1$ be arbitrary and observe that

$$(x \ast |a|)_k = \sum_{k_1=k-N_i+1}^{k+N_i-1} x_{k_1} |a|_{k-k_1},$$

since $a_{k_2} = 0$ for $k_2 \geq N_i$. Therefore,

$$\varphi^-_a (\varepsilon_l) = \begin{cases} 
    |a_{N_i-1}| \frac{\nu^{N_i-1}_i}{N_i}, & l = 0, \\
    \frac{1}{2} \sum_{k=N_i-1-l}^{N_i-1} |a|_k \frac{\nu^k_i}{k+l+1}, & 1 \leq l \leq 2 (N_i - 1), \\
    \frac{1}{2} \sum_{k=1-N_i}^{N_i-1} |a|_k \frac{\nu^k_i}{k+l+1}, & l \geq 2 (N_i - 1). 
\end{cases}$$

61
In particular, note that \( \varphi_a^- (\varepsilon_l) \) is decreasing for \( l \geq 2 (N_i - 1) \). Hence

\[
\| \varphi_a^- \|_{\nu_i}^* = \sup_{l \in \mathbb{N}_0} | \varphi_a^- (\varepsilon_l) | = \frac{1}{2} \Gamma_a^-
\]

by Lemma 2.2.5.

An analogous computation yields \( \| \varphi_a^+ \|_{\nu_i}^* = \frac{1}{2} \Gamma_a^+ \).

\[\square\]

**Corollary 2.6.11** (Estimates for the tail). Define

\[
d_1^i : = \frac{\nu_i^2 + 1}{\nu_i N_i} \sum_{j=1}^n | g_{e_j} | + \sum_{1 \leq l \leq s \leq n} | g_{l_s} | \left( \nu_i \left( \Gamma_{\hat{a}_i}^- + \Gamma_{\hat{a}_i}^- \right) + \frac{1}{\nu_i} \left( \Gamma_{\hat{a}_i}^+ + \Gamma_{\hat{a}_i}^+ \right) \right),
\]

\[
d_2^i : = \frac{2 (\nu_i^2 + 1)}{\nu_i N_i} \sum_{|\alpha|=2} | g_\alpha |,
\]

then

\[
2 \sum_{k=N_i}^\infty \left| \frac{d}{d\tau} \right|_{\tau=0} \left( c_{k-1} (\hat{a}_i + rv + \tau w) - c_{k+1} (\hat{a}_i + rv + \tau w) \right) \frac{\nu_i^k}{k} \leq d_1^i + rd_2^i.
\]

**Proof.** First note that

\[
2 \sum_{k=N_i}^\infty \left| \frac{d}{d\tau} \right|_{\tau=0} c_{k-1} (\hat{a}_i + rv + \tau w) \frac{\nu_i^k}{k} = 2\nu_i \sum_{k=N_i-1}^\infty \left| \frac{d}{d\tau} \right|_{\tau=0} c_k (\hat{a}_i + rv + \tau w) \frac{\nu_i^k}{k+1}.
\]

Next, observe that

\[
2 \sum_{k=N_i-1}^\infty \left| \sum_{j=1}^n g_{e_j} [w_k]_j \right| \frac{\nu_i^k}{k+1} \leq \frac{1}{N_i} \sum_{j=1}^n | g_{e_j} | \left\| [w]_j \right\|_{\nu_i} \leq \frac{1}{N_i} \sum_{j=1}^n | g_{e_j} |,
\]

while

\[
2 \sum_{k=N_i-1}^\infty \left| (w_l * \hat{a}_s^i) \right|_k \frac{\nu_i^k}{k+1} \leq 2 \varphi_{\hat{a}_s^-} (|w_l|) \leq \Gamma_{\hat{a}_s^-},
\]

and

\[
2 \sum_{k=N_i-1}^\infty \left| (w_l * rv_s) \right|_k \frac{\nu_i^k}{k+1} \leq \frac{r}{N_i} \left\| w_l * v_s \right\|_{\nu_i} \leq \frac{r}{N_i},
\]

62
for any \( l, s \in \{1, \ldots, n\} \) by Proposition 2.2.4, Lemma 2.6.10 and since \( v, w \in B_1(0) \). Therefore,

\[
2 \sum_{k=N_i}^{\infty} \left| \frac{d}{d\tau} \right|_{\tau=0} c_{k-1} \left( \hat{\alpha}^i + rv + \tau w \right) \left\lfloor \frac{\nu^i_k}{k} \right\rfloor \\
\leq \nu_i \left( \frac{1}{N_i} \sum_{j=1}^{n} |g_{e_j}| + \sum_{1 \leq l \leq s \leq n} |g_{ls}| \left( \Gamma^-_{\hat{\alpha}_i^l} + \Gamma^-_{\hat{\alpha}_i^s} + \frac{2r}{N_i} \right) \right),
\]

by (2.39). An analogous computation shows that

\[
2 \sum_{k=N_i}^{\infty} \left| \frac{d}{d\tau} \right|_{\tau=0} c_{k+1} \left( \hat{\alpha}^i + rv + \tau w \right) \left\lfloor \frac{\nu^i_{k+1}}{k} \right\rfloor \\
\leq \frac{1}{\nu_i} \left( \frac{1}{N_i} \sum_{j=1}^{n} |g_{e_j}| + \sum_{1 \leq l \leq s \leq n} |g_{ls}| \left( \Gamma^+_{\hat{\alpha}_i^l} + \Gamma^+_{\hat{\alpha}_i^s} + \frac{2r}{N_i} \right) \right),
\]

which completes the proof of the lemma.

We are now ready to define the \( Z \)-bounds:

**Proposition 2.6.12.** The bounds

\[
Z_{i,j}(r) := \left[ \gamma \| \Pi_{i,j}A_N \|_{B(\mathcal{X}_{\nu_i}^{\infty}, t_{\nu_i})} + \frac{1}{\omega^2} \left( t_i - t_{i-1} - \frac{1}{4} \left| d_j^i \right| + 2 \right) \right] r^2 \\
+ \left[ h_{i,j} + \| \Pi_{i,j}(Z_1) \|_{\nu_i} + \frac{t_i - t_{i-1}}{4\omega} \left| d_j^i \right| \right] r,
\]

\[
Z_0(r) := \gamma \| \Pi_0A_N \|_{B(\mathcal{X}_{\nu_0}^{\infty}, \mathcal{R})} r^2 + (h_0 + \Pi_0(Z_1)) r,
\]

where \( 1 \leq i \leq m, 1 \leq j \leq n \), satisfy (2.13) and (2.15), respectively.

**Proof.** First observe that

\[
\left( \Pi_iA \left( DF(\hat{x} + w_1) - \hat{A} \right) w \right)_k \\
= - \frac{t_i - t_{i-1}}{4\omega k} \left| \frac{d}{d\tau} \right|_{\tau=0} \left( c_{k-1} \left( \hat{\alpha}^i + rv + \tau w \right) - c_{k+1} \left( \hat{\alpha}^i + rv + \tau w \right) \right) \\
+ \frac{\omega w v^i_k + \omega v w^i_k}{\omega} r
\]

for all \( k \geq N_i \) by (2.28) and (2.31) and

\[
\sum_{k=N_i}^{\infty} \left| \omega w v^i_k + \omega v w^i_k \right| \nu^i_k \leq 1_n, \quad (2.40)
\]
since \(v, w \in B_1(0)\). Now recall the decomposition in (2.11) and combine (2.40), Lemma 2.6.4, Proposition 2.6.8, and Corollary 2.6.11 to obtain the result.

### 2.6.3 Modifications for non-periodic boundary conditions

In Section 2.5.3 an equation representing arc-length continuation is added to the system, accompanied by an extra a priori unknown parameter. In Section 2.5.4 the periodic boundary conditions are replaced by boundary conditions that guarantee that the solution ends up in the local stable and unstable manifolds. The adaptations of the estimates to these modified problems are presented in Sections 2.6.3 and 2.6.3, respectively.

**Periodic solutions near the homoclinic orbit**

In this section we incorporate the necessary adjustments for the modified problem introduced in section 2.5.3. First, observe that we will have an additional radii-polynomial \(p_{-1}\) which corresponds to the equation for \(\rho\). In particular, the additional bound \(Y_{-1}\) is given by

\[
Y_{-1} := |\Pi_{-1} A_N F_N(\hat{x})|,
\]

and the formulae for the other \(Y\)-bounds remain the same.

Recall that the \(Z\)-bounds were derived by estimating the two terms in (2.11). In particular, one can derive bounds for the norm of linear operators on \(X^N_\nu\) in exactly the same way as before as in Lemma 2.6.2 and then use Lemma 2.6.4 to compute a bound for \((I - A\hat{A})x_2\). The changes in the bounds for the second term, \(A\left(DF(\hat{x} + x_1) - \hat{A}\right)x_2\), are more subtle, since \(\rho\) is now to be interpreted as an unknown variable as well.

To identify the differences, denote the approximate solution by \(\hat{x} = (\hat{\rho}, \hat{\omega}, \hat{a}^1, \ldots, \hat{a}^m)\), write \(x_1 = rv, x_2 = rw\), where \(v = (\rho_1, \omega_1, v^1, \ldots, v^m), w = (\rho_2, \omega_2, w^1, \ldots, w^m) \in B_1(0)\), and recall that the bounds for \(A\left(DF(\hat{x} + x_1) - \hat{A}\right)x_2\) were obtained by computing estimates for (2.30) and (2.31). Furthermore, observe that the additional equation for \(\rho\) has no contribution to this part of the analysis, since the equation is linear and therefore

\[
\left. \frac{d}{d\tau} \right|_{\tau=0} \Pi_{-1} (F(\hat{x} + rv + \tau w) - F_N(\hat{x} + \tau w)) = 0.
\]

Next, note that the convolution terms \(c^i\) are the only functions in the definition of \(F\) which depend on \(\rho\), since \(g_1\) is the only coefficient in the Lorenz-system which depends on \(\rho\). Consequently, a straightforward computation shows that the term

\[
\begin{bmatrix}
0 \\
0 \\
r (\rho_1 [w^i]_1 + \rho_2 [v^i]_1)
\end{bmatrix},
\]

64
needs to be added to the right-hand side of (2.32), and \( g_{c_1} \left[ \tilde{w}^i \right]_1 = g_{c_1} (\hat{\rho}) \left[ \tilde{w}^i \right]_1 \).

To incorporate this extra term in the estimates for (2.35) one needs to modify Proposition 2.6.8 by setting 

\[
\gamma := \max \left\{ 2 (N_i - 1) + \frac{(t_i - t_{i-1}) (2\nu_i^2 + 1)}{2\nu_i} \left( \sum_{|\alpha|=2} |g_{\alpha}| + \left[ \begin{array}{c} 0 \\ 1 \end{array} \right] \right) \right\}^m_{i=1} \right\}.
\]

Similarly, terms

\[
\begin{bmatrix} 0 \\ \rho_2 \left[ \hat{a}^i \right]_1 \end{bmatrix} + \begin{bmatrix} 0 \\ \rho_2 \left[ \nu^i \right]_1 + \rho_1 \left[ w^i \right]_1 \end{bmatrix} \text{r},
\]

need to be added to the right-hand side of (2.39). Therefore, by analogous computations as in the proof of Corollary 2.6.11 the tail-estimates in Corollary 2.6.11 remain valid if we add the terms

\[
2\nu_i^N \left[ \left[ \hat{a}^i_{N_i-1} \right]_1 \right]_{N_i} \quad \text{and} \quad 2 \left( \nu_i^2 + 1 \right) \left[ \left[ \nu^i \right]_1 \right]_{N_i} \nu_i,
\]

to the expressions for \([d^i_1]\) and \([d^i_2]\), respectively.

The above modifications account for all the necessary changes in the estimates, and the extra formula for the additional bound \( Z_{-1} \) is given by

\[
Z_{-1}(r) := \gamma \| \Pi_{-1} A_N \|_{\mathcal{B}(X^N_{\nu}, \mathbb{R})} r^2 + (h_{-1} + \Pi_{-1} (Z_1)) r.
\]

**Heteroclinic orbit**

In this section we identify the differences in the construction of the bounds for the connecting orbit discussed in Section 2.7. We start by identifying the differences in the \( Y \)-bounds. We write \( \hat{x} = \left( \hat{\phi}, \hat{\alpha}, \hat{a}^1, \ldots, \hat{a}^m \right) \) and observe that the main difference in the computation of the \( Y \)-bounds, as performed in Proposition 2.6.1, is caused by the following two terms being nonzero:

\[
\left[ \Pi_{0,j} (F (\hat{x}) - F_N (\hat{x})) \right]_{j=1}^3 = (Q - Q_{N_s}) \circ \Theta_{\mu} (\hat{\phi}),
\]

and

\[
\left( \Pi_1 (F (\hat{x}) - F_N (\hat{x})) \right)_0 = (P - P_{N_u}) (\hat{\alpha}).
\]

Consequently, we have additional bounds corresponding to the finite dimensional part of \( A(F (\hat{x}) - F_N (\hat{x})) \), which previously had no contribution at all.

The required modifications are as follows: set

\[
\delta := |A_N| \left( \begin{array}{c} \delta_s \cdot 1_n \\ \delta_u \cdot 1_n \\ 0_n (-1 + \sum_{i=1}^m N_i) \end{array} \right),
\]
add the bounds $\|\Pi_{i,j} (\delta)\|_{\nu_i}$ to $Y_{i,j}$ in Proposition 2.6.1, define (instead of $Y_0$) the bounds

$$Y_{0,j} := |\Pi_{0,j} A_N F_N (\hat{x})| + \Pi_{0,j} (\delta),$$

where $1 \leq j \leq n$, and change the factor $(\hat{\omega})^{-1}$ into $L$.

Next, we consider the computation of the $Z$-bounds by considering the decomposition in (2.11) again. As before, the main differences occur in the bounds for $A \left( DF (\hat{x} + x_1) - \hat{A} \right) x_2$, which stem from the fact that there is no dependence on $\omega$ anymore, whereas dependencies on $\phi$ and $\alpha$ need to be incorporated. To identify the differences, write $x_1 = rv$ and $x_2 = rw$, where $v = (\phi_1, \alpha_1, v^1, \ldots, v^m)$, $w = (\phi_2, \alpha_2, w^1, \ldots, w^m) \in B_1 (0)$.

The main differences in the right-hand sides of (2.30) and (2.31) in the current setting are that the term $k (\omega_2 v^i_k + \omega_1 w^i_k) r$ is not present, and the factors $t_i - t_{i-1}$ need to be multiplied by $L$. To incorporate these changes into the bounds for (2.35) we only need to modify Proposition 2.6.8 by setting

$$\gamma := \max \left( \{ |\Lambda_{0,2}|_{\infty} \} \cup \left\{ \left| \Lambda_{1,2} + \frac{L (t_1 - t_0) (2\nu_1^2 + 1)}{2\nu_1} \sum_{|\alpha|=2} |g_\alpha| \right|_{\infty} \right\} \right)$$

$$\cup \left\{ \left| \frac{L (t_i - t_{i-1}) (2\nu_i^2 + 1)}{2\nu_i} \sum_{|\alpha|=2} |g_\alpha| \right|_{\infty} : 2 \leq i \leq m \right\} \right).$$

Here the extra bounds $\Lambda_{0,1}, \Lambda_{0,2} \in \mathbb{R}^2_+$ and $\Lambda_{1,1}, \Lambda_{1,2} \in \mathbb{R}^3$, defined analogously as in (2.36) and (2.37), respectively, can be obtained by computing estimates for

$$\left. \frac{d}{d\tau} \right|_{\tau=0} \Pi_{0,j} (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w)) = 2 \sum_{k=N_m}^{\infty} [w^m_k]_j$$

$$- \left. \frac{d}{d\tau} \right|_{\tau=0} \left( Q_{N_s} \circ \Theta_{\mu} \left( \hat{\phi} + r \phi_1 + \tau \phi_2 \right) - Q_{N_s} \circ \Theta_{\mu} \left( \hat{\phi} + \tau \phi_2 \right) \right)$$

$$+ h_s \circ \Theta_{\mu} \left( \hat{\phi} + r \phi_1 + \tau \phi_2 \right)$$

for $1 \leq j \leq 3$, and

$$\left. \frac{d}{d\tau} \right|_{\tau=0} \left( \Pi_1 (F (\hat{x} + rv + \tau w) - F_N (\hat{x} + \tau w)) \right)_0 = 2 \sum_{k=N_1}^{\infty} (-1)^k w^1_k$$

$$- \left. \frac{d}{d\tau} \right|_{\tau=0} \left( P_{N_u} (\hat{\alpha} + r \alpha_1 + \tau \alpha_2) - P_{N_u} (\hat{\alpha} + \tau \alpha_2) + h_u (\hat{\alpha} + r \alpha_1 + \tau \alpha_2) \right),$$

respectively, where $h_u = P - P_{N_u}$ and $h_s = Q - Q_{N_s}$. The series involving $w^m$ and $w^1$ in the latter two expressions can be bounded in the same way as in Lemma
The terms associated to the parameterizations \( P \) and \( Q \) can be bounded by using a combination of analysis and interval arithmetic. In particular, this involves the choice of an a priori radius \( r^* > 0 \) in order to compute uniform bounds for \(|DP_{N_u} (\alpha^* + r\alpha_1)|\) and \(|D (Q_{N_s} \circ \Theta_\mu)(\phi^* + r\phi_1)|\) for \( 0 < r < r^* \), where \(|\phi_1|, |\alpha_1|_2 \leq 1\). The reader is referred to [49] for the details.

The final changes to be made are in Proposition 2.6.12: the \( Z \)-bounds are now defined by

\[
Z_{i,j} (r) := \left[ \gamma \left\| \Pi_{i,j} A_N \right\|_{\mathcal{B}(\mathcal{X}_{\nu_i}, \mathcal{E}_{i_1})} + \frac{L(t_i - t_{i-1})}{4} \left[ d_{i_1}^2 \right]_{j} \right] r^2 \\
+ \left[ h_{i,j} + \left\| \Pi_{i,j} (Z_1) \right\|_{\nu_i} + \frac{L(t_i - t_{i-1})}{4} \left[ d_{i_1}^2 \right]_{j} \right] r,
\]

\[
Z_{0,j} (r) := \gamma \left\| \Pi_{0,j} A_N \right\|_{\mathcal{B}(\mathcal{X}_{\nu}, \mathcal{E})} r^2 + (h_{0,j} + \Pi_{0,j} (Z_1)) r,
\]

where \( 1 \leq i \leq m \) and \( 1 \leq j \leq 3 \).
Validated computations for connecting orbits in polynomial vector fields

3.1 Introduction

Connecting orbits play a central role in the study of dynamical systems. They provide a detailed picture of how a dynamical system can evolve from one “state” (e.g. an equilibrium, a periodic orbit or another type of recurrent set) into another. Furthermore, their existence can often be used to establish more complicated dynamical phenomena through forcing theorems. However, proving the existence of a connecting orbit for a given nonlinear ODE is in general a difficult (if not impossible) task to accomplish by hand. For this reason, one often resorts to numerical methods. While numerical methods can provide valuable insight into quantitative properties of a connecting orbit, which would otherwise be out of reach with merely a pen and paper analysis, the results are usually non-rigorous and cannot be used in mathematical arguments. In particular, a standard numerical method does not yield a proof for the existence of a connecting orbit.

In this paper we present a general computer-assisted method for proving the existence of transverse connecting orbits between hyperbolic equilibria for nonlinear ODEs. The method is based on solving the finite time boundary value problem

\[
\begin{align*}
\frac{du}{dt} &= g(u), & t &\in [0, L], \\
u(0) &\in W^u_{\text{loc}}(p_0), \\
u(L) &\in W^s_{\text{loc}}(q_0),
\end{align*}
\]

(3.1)

where \( g : \mathbb{R}^n \to \mathbb{R}^n \) is a general polynomial vector field, \( p_0, q_0 \in \mathbb{R}^n \) are hyperbolic equilibria and \( L > 0 \) is the time needed to travel between the local (un)stable manifolds. We assume that \( \dim(W^u(p_0)) + \dim(W^s(q_0)) = n + 1 \), which is a necessary condition for a transverse connecting orbit to exist. The idea is to solve (3.1) by computing Taylor expansions for charts on the local (un)stable manifolds via the parameterization method [14,75], which are used to supplant the boundary...
conditions in (3.1) with explicit equations, and to use the domain decomposition techniques based on Chebyshev series developed in Chapter 2 to parameterize the orbit in between. We remark that the assumption that $g$ is polynomial is not as restrictive as it initially might seem, since many nonlinearities which consist of elementary functions can be brought into polynomial form by using automatic differentiation techniques, see [48] for instance.

Before we proceed with a more detailed description of our method, a few remarks concerning the development of numerical methods for connecting orbits are in order. Many numerical methods (both rigorous and non-rigorous) are based on approximating (un)stable manifolds and solving finite time boundary value problems. We mention the (non-rigorous) methods implemented in the continuation packages MATCONT [28] and AUTO [1] in particular. Furthermore, many validated numerical methods have been developed over the last decade. It is beyond the scope of this text to give an overview. Nevertheless, we mention the functional analytic methods developed in [5, 8, 46, 49, 50, 72, 74, 76], which are based on solving fixed point problems, the topological methods developed in [80–83], which are based on covering relations, cone conditions [41, 89, 91] and rigorous integration of the flow via Lohner-type algorithms [87], and the methods in [21, 44] based on shadowing techniques.

The first step in the development of our validated numerical method is to recast (3.1) into an equivalent zero finding problem $F(x) = 0$. The unknowns in this problem are the Taylor coefficients of the parameterizations of the local (un)stable manifolds, which include the equilibria and the associated eigendata, the coordinates of the endpoints $u(0)$ and $u(L)$ on the associated charts, and the Chebyshev coefficients of the orbit. Next, we use the computer to determine an approximate zero of $F$. The numerical computations are then combined with analysis on paper to construct a Newton-like map $T$ whose fixed points correspond to zeros of $F$. Finally, we use pen and paper estimates to derive a finite number of inequalities, which can be used to determine a neighborhood around the approximate solution on which $T$ is a contraction. An essential property of these inequalities is that they can be rigorously verified with the aid of a computer. This approach is in the literature often referred to as a parameterized Newton-Kantorovich method or the radii-polynomial approach (see [27, 84]).

The methods developed in [50, 72, 76] are very similar to ours; they also use the parameterization method, Chebyshev series and the radii-polynomial approach to validate solutions of (3.1). The main difference is that in [50, 72, 76] the charts on the local (un)stable manifolds and the connecting orbit in between are validated separately. More precisely, in these papers the strategy is to first validate the charts on the local (un)stable manifolds, which amounts to solving two separate zero finding problems, and to validate the connecting orbit afterwards (which constitutes a third zero finding problem). The novelty of our method is that the charts on the local (un)stable manifolds and connecting orbit in between are validated simultaneously as the locally unique zero of one map. This makes the proposed method well-suited for rigorous continuation and bifurcation studies of connecting orbits. In addition, while the methods in [50, 72, 76] are in principle applicable to a general
class of equations, the code and estimates were only developed in detail for specific problems. In this paper, the mathematical analysis and the code are developed in full detail for general polynomial vector fields. In particular, the presented method deals in a universal and systematic manner with the cases of real and complex eigenvalues associated to the parameterizations of the (un)stable manifolds (see the examples below).

Before we present some applications, let us discuss a few possible extensions. The first extension is to let the vector field explicitly depend on a parameter and to perform rigorous (pseudo-arclength) continuation of connecting orbits. This involves a relatively straightforward application of the uniform contraction principle and a slight modification of the estimates developed in this paper (see [15,73,76] for instance). Furthermore, in order to carry out continuation efficiently, we need to develop algorithms (heuristics) which automatically determine near-optimal parameter values for the validation of the charts on the local (un)stable manifolds and the connecting orbit. More specifically, during continuation it might become necessary to modify the number of Taylor coefficients, the size of the charts on the local (un)stable manifolds, the grid on which the connecting orbit is computed, the number of Chebyshev coefficients, or the integration time.

The second extension involves the incorporation of resonances. In this paper, we assume that the (un)stable eigenvalues associated to the (un)stable manifolds satisfy a so-called non-resonance condition. This condition is related to the regularity of the chart mappings obtained via the parameterization method. In short, the parameterization method is based on constructing a smooth conjugacy (analytic in our case) between the nonlinear flow on the (un)stable manifold and an “easier” fully understood model system. One can choose this model system to be linear, which we do in this paper, if the (un)stable eigenvalues satisfy a non-resonance condition. If there are resonant eigenvalues, however, one needs to use a nonlinear model system instead. This is explained in detail in [75]. Generically, one will encounter resonances during continuation. Therefore, in order to successfully perform validated continuation, we need to allow for the possibility of resonant eigenvalues and modify the current method accordingly as explained in [75]. In particular, we need to develop an algorithm which automatically detects when to “switch” between the linear and nonlinear model flow during continuation. Furthermore, a careful analysis of the case in which a pair of complex conjugate eigenvalues become real (or vice versa) is needed as well. After the above extensions have been implemented, one can start developing tools for the rigorous study of bifurcations of connecting orbits for nonlinear ODEs.

The computer-assisted method presented in this paper is implemented in an object oriented framework in MATLAB using the INTLAB package [66] for interval arithmetic. The code is available at [67]. A third and useful extension would be to incorporate an extra degree of freedom into the classes for the connecting orbit so that additional equations and variables can be added (or removed) in a convenient manner. This would facilitate the required modifications for dealing with non-polynomial vector fields via automatic differentiation techniques, analyzing connecting orbits in vector fields with symmetry, proving the existence of
Figure 3.1: Validated traveling wave profiles of \((3.2)\) for \(a = 5\), \(D = 3\), \(b = \frac{1}{2}\) and \(\kappa = -0.7767\). The depicted parts of the traveling wave profiles correspond to the first and third component of the connecting orbit between the unstable and stable manifold of \((b, 0, 1 - b, 0)\) and \((1, 0, 0, 0)\), respectively. The time of flight of the connection between the (un)stable manifolds was \(L = 20\).

homoclinic instead of heteroclinic orbits, and performing bifurcation analysis.

Finally, we remark that the computational efficiency of the current implementation can be improved. For instance, the equations for the parameterizations of the local (un)stable manifolds and the connecting orbit in between are to a large extent uncoupled. As a consequence, the derivative of the zero finding map \(F\) has a block structure, which can be exploited to reduce the computational costs of the computation of an approximate inverse (we need an explicit finite dimensional approximate inverse to construct a Newton-like map \(T\)). Furthermore, in applications it might not be necessary to resolve the full local stable manifold equally well in all directions, but a rather more focussed parameterization centered around the slow eigendirections of the equilibria is appropriate, since a connecting orbit generically tends to enter the stable manifold via these directions.

Application 3.1.1 (Traveling fronts in the Lotka-Volterra equations). The Lotka-Volterra equations are a system of reaction-diffusion equations given by

\[
\begin{align*}
\frac{\partial v}{\partial t} &= D \frac{\partial^2 v}{\partial x^2} + v (1 - v - w), \\
\frac{\partial w}{\partial t} &= \frac{\partial^2 w}{\partial x^2} + aw (v - b),
\end{align*}
\tag{3.2}
\]

Application 3.1.1
where $D > 0$, $a > 0$, $b \in (0, 1)$ and $(t, x) \in \mathbb{R}^2$. This system has three homogeneous equilibrium states: $(v, w) = (0, 0)$, $(v, w) = (1, 0)$ and $(v, w) = (b, 1 - b)$. We have used our method to prove the existence of solutions of (3.2) of the form $v(t, x) = \zeta_1(x - \kappa t)$ and $w(t, x) = \zeta_2(x - \kappa t)$, where $\zeta_1, \zeta_2 : \mathbb{R} \to \mathbb{R}$ and $\kappa < 0$, which satisfy

$$\lim_{\tau \to -\infty} (\zeta_1 (\tau), \zeta_2 (\tau)) = (1, 0), \quad \lim_{\tau \to \infty} (\zeta_1 (\tau), \zeta_2 (\tau)) = (b, 1 - b).$$

Such solutions are often referred to as \textit{traveling fronts} with wave speed $\kappa$.

Substitution of the traveling wave Ansatz $(\zeta_1, \zeta_2)$ into (3.2) shows that connecting orbits from $(b, 0, 1 - b, 0)$ to $(1, 0, 0, 0)$ for the four dimensional system of ODEs

$$\frac{du}{dt} = \begin{bmatrix} -u_2 \\ D^{-1} (\kappa u_2 + u_1 (1 - u_1 - u_3)) \\ -u_4 \\ (\kappa u_4 + au_3 (u_1 - b)) \end{bmatrix} \quad (3.3)$$

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.2.png}
\caption{A three dimensional projection of a validated connecting orbit of (3.3) for $a = 5$, $D = 3$, $b = \frac{1}{2}$ and $\kappa = -0.7767$. The geometric objects colored in red and blue correspond to parameterizations of the local unstable and local stable manifold of $(b, 0, 1 - b, 0)$ and $(1, 0, 0, 0)$, respectively, which were computed (and validated) by using the parameterization method. The curve in black corresponds to the piece of the connecting orbit which was validated by using Chebyshev series. The time of flight of the connection between the (un)stable manifolds was $L = 20$. We remark that the integration time was not optimized, i.e., it is possible to decrease $L$ and to “absorb” more of the orbit into the parameterized local (un)stable manifolds.}
\end{figure}
Figure 3.3: A validated traveling wave profile of (3.4) for $a = -0.1$, $\kappa = -2$ and $\gamma = 4.202$. The depicted part of the traveling wave profile corresponds to the first component of the connecting orbit between the unstable and stable manifold of $(-1, 0, 0, 0)$ and $(a, 0, 0, 0)$, respectively. The time of flight of the connection between the (un)stable manifolds was $L = 4$.

correspond to traveling wave profiles $(\zeta_1(t), \zeta_2(t)) = (u_1(-t), u_3(-t))$ and vice versa. We have successfully validated connecting orbits in (3.3) for various values of $\kappa \in [-1, -0.5938]$. For these parameter values, the equilibria $(b, 0, 1 - b, 0)$ and $(1, 0, 0, 0)$ have a two dimensional unstable and three dimensional stable manifold, respectively. In particular, the stable eigenvalues of the linearization at $(1, 0, 0, 0)$ consist of one complex conjugate pair of eigenvalues and one real eigenvalue. We have depicted a validated traveling wave profile and the corresponding connecting orbit for a particular wave speed in Figures 3.1 and 3.2, respectively. The reader is referred to Section 3.6.1 for the details.

**Application 3.1.2 (Traveling fronts in a fourth order parabolic PDE).** We have proven the existence of traveling fronts $v(t, x) = \zeta(x - \kappa t)$ for the following fourth order parabolic PDE:

$$
\frac{\partial v}{\partial t} = -\gamma \frac{\partial^4 v}{\partial x^4} + \frac{\partial^2 v}{\partial x^2} + (v - a)(1 - v^2)
$$

(3.4)

where $-1 < a \leq 0$ and $\gamma > 0$. The parameter values were taken from [4] in which geometric singular perturbation theory was used to prove the existence of traveling fronts between the homogeneous states $v \equiv -1$ and $v \equiv 1$ for sufficiently small
Figure 3.4: A three dimensional projection of a validated connecting orbit of (3.5) for $a = -0.1$, $\kappa = -2$ and $\gamma = 4.202$. The geometric objects colored in red and blue correspond to parameterizations of the local unstable and local stable manifold of $(-1,0,0,0)$ and $(a,0,0,0)$, respectively, which were computed (and validated) by using the parameterization method. The curve in black corresponds to the piece of the connecting orbit which was validated by using Chebyshev series. The connecting orbit was relatively “short” and the time of flight was $L = 4$. We remark that the integration time was not optimized.

We have successfully established the existence of connecting orbits from $(-1,0,0,0)$ to $(a,0,0,0)$ for the four dimensional system of ODEs

$$\frac{du}{dt} = -\begin{bmatrix} \gamma u_2 \\ \gamma u_3 \\ \gamma u_4 \\ \kappa u_2 + u_3 + (u_1 - a) (1 - u_1^2) \end{bmatrix},$$

which correspond to traveling wave profiles $\zeta(t) = u_1 \left( -\frac{t}{\gamma} \right)$, for a fixed wave speed $\kappa$ and various values of $\gamma \in [0.4557, 10.50]$. We rescaled time with a factor $\gamma$ so that the system in (3.5) is well-defined at $\gamma = 0$. We have depicted a validated traveling wave profile and the corresponding connecting orbit in Figures 3.3 and 3.4 respectively. The reader is referred to Section 3.6.2 for the details.

This paper is organized as follows. In Section 4.2, we review some basic facts about Chebyshev series, Taylor series and sequence spaces, which will be used
extensively throughout this paper. In Section 3.3 we set up an equivalent zero finding problem for (3.1) by using domain decomposition, the parameterization method, Chebyshev series and Taylor series. In Section 3.4 we set up an equivalent fixed-point problem and explain how the existence of a zero can be established with the aid of a computer. This involves the construction of computable bounds which are developed in full detail in Section 3.5. Finally, in Section 4.9 we demonstrate the effectiveness of the method by proving the existence of traveling fronts in parabolic PDEs.

3.2 Preliminaries

In this section we develop a functional analytic framework for analyzing maps which arise from the study of connecting orbits. We start in Section 3.2.1 by recalling basic results from Chebyshev approximation theory. In Sections 3.2.2 and 3.2.3 we introduce spaces of geometrically decaying sequences and multivariate arrays, respectively. In addition, we review methods for analyzing bounded linear operators on them.

3.2.1 Chebyshev series

In this section we recall basic notions and results from Chebyshev approximation theory. The reader is referred to [68] for the proofs and a more comprehensive introduction into the theory of Chebyshev approximations.

**Definition 3.2.1.** The Chebyshev polynomials $T_k : [-1, 1] \to \mathbb{R}$ are defined by the relation $T_k (\cos (\theta)) = \cos (k \theta)$, where $k \in \mathbb{N}_0$ and $\theta \in [0, \pi]$.

Chebyshev series constitute a non-periodic analog of Fourier cosine series and have similar convergence properties. For instance, any Lipschitz continuous function admits a unique Chebyshev expansion. In this paper, we will consider Chebyshev expansions of analytic functions. The Chebyshev coefficients of such regular functions decay (in analogy with Fourier series) at a geometric rate to zero. A more precise statement is given in the next proposition.

**Proposition 3.2.2.** Suppose $u : [-1, 1] \to \mathbb{R}$ is analytic and let

$$u = a_0 + 2 \sum_{k=1}^{\infty} a_k T_k$$

be its Chebyshev expansion. Let $E_\nu \subset \mathbb{C}$ denote an open ellipse with foci $\pm 1$ to which $u$ can be analytically extended, where $\nu > 1$ is the sum of the semi-major and semi-minor axis of $E_\nu$. If $u$ is bounded on $E_\nu$, then $|a_k| \leq M \nu^{-k}$ for all $k \in \mathbb{N}_0$, where $M = \sup_{z \in E_\nu} |u(z)|$.

The Chebyshev coefficients of the product of two Chebyshev series is (in direct analogy with Fourier cosine series) given by the symmetric discrete convolution:
Proposition 3.2.3. Suppose \( u, v : [-1, 1] \to \mathbb{R} \) are Lipschitz continuous and let

\[
u = a_0 + 2 \sum_{k=1}^{\infty} a_k T_k, \quad v = b_0 + 2 \sum_{k=1}^{\infty} b_k T_k,
\]

be the associated Chebyshev expansions. Then

\[
u \cdot v = (a \ast b)_0 + 2 \sum_{k=1}^{\infty} (a \ast b)_k T_k, \quad \text{where} \quad (a \ast b)_k := \sum_{k_1 + k_2 = k} a_{k_1} b_{k_2}.
\]

3.2.2 Geometrically decaying sequences

In this section we introduce a sequence space suitable for analyzing analytic functions, and elementary operations on them, via their Chebyshev coefficients. Recall that the Chebyshev coefficients of an analytic function \( u : [a, b] \to \mathbb{R}^n \) decay exponentially fast to zero by Proposition 3.2.2. In light of this observation we define

\[
\ell^1_{\nu, n} := \left\{ (a_k)_{k \in \mathbb{N}_0} : a_k \in \mathbb{C}^n, \sum_{k=0}^{\infty} |a_k|_j \nu^k < \infty, 1 \leq j \leq n \right\},
\]

where \([a_k]_j\) denotes the \( j \)-th component of \( a_k \) and \( \nu > 1 \) is some prescribed weight, endowed with the norm

\[
\|a\|_{\nu, n} := \max_{1 \leq j \leq n} \left\{ |a_0|_j + 2 \sum_{k=1}^{\infty} |a_k|_j \nu^k \right\}.
\]

In the special case that \( n = 1 \) we shall write \( \ell^1_{\nu} := \ell^1_{\nu, 1} \) and \( \|\cdot\|_{\nu} := \|\cdot\|_{\nu, 1} \). It is a straightforward task to verify that \( \ell^1_{\nu, n} \) equipped with this norm is a Banach space over \( \mathbb{C} \).

Remark 3.2.4. In this paper we are exclusively concerned with Chebyshev expansions of real-valued functions. From this perspective it is more natural to consider sequence spaces over \( \mathbb{R} \) instead of \( \mathbb{C} \). The reason for using a space of complex valued sequences is that we wish to couple the Chebyshev expansions with chart maps for (un)stable manifolds, which might be complex-valued (see Section 3.3). We will proof a-posteriori that the Chebyshev coefficients are real by using arguments based on symmetry.

The operation of multiplying two Chebyshev series can be lifted to the level of sequences, giving rise to the symmetric discrete convolution \( \ast \), as shown in Proposition 3.2.3. This additional product structure on \( \ell^1_{\nu} \) yields a particularly nice space:

Proposition 3.2.5. The space \((\ell^1_{\nu}, \ast)\) is a commutative Banach algebra over \( \mathbb{C} \).

Proof. This follows directly from Proposition 3.2.3 and the triangle inequality. \( \square \)
One of the reasons for using the space $\ell^1_\nu$ is to have a relatively simple and sharp convolution estimate. Another important reason is that it is easy to compute the norm of bounded linear operators. To explain how to compute the norm of a bounded linear operator on $\ell^1_\nu$, we introduce the notion of the corner points. Let $\{e_k\}_{k\in\mathbb{N}_0}$ denote the canonical Schauder basis for $\ell^1_\nu$, i.e. $(e_k)_l := \delta_{kl}$ for $l \in \mathbb{N}_0$, so that

$$a = \sum_{k=0}^{\infty} a_k e_k,$$

for any $a \in \ell^1_\nu$.

**Remark 3.2.6.** We shall frequently use the Schauder basis $(e_k)_{k\in\mathbb{N}_0}$ to identify an element $a \in \ell^1_\nu$ with the infinite dimensional column vector $[a_0 \ a_1 \ \ldots]^T$.

**Definition 3.2.7.** The corner points $\{\xi_{k,\nu}\}_{k\in\mathbb{N}_0} \subset \ell^1_\nu$ of the unit ball in $\ell^1_\nu$ are defined by $\xi_{k,\nu} := \varepsilon_{k,\nu} e_k$, where

$$\varepsilon_{k,\nu} = \begin{cases} 1 & k = 0, \\ \frac{1}{2}^{\nu-k}, & k \in \mathbb{N}. \end{cases}$$

We shall write $\xi_{k,\nu} = \xi_k$ and $\varepsilon_{k,\nu} = \varepsilon_k$ whenever there is no chance of confusion.

The norm of a bounded linear operator on $\ell^1_\nu$ can be computed by simply evaluating it at the corner points as shown in the next proposition:

**Proposition 3.2.8.** Let $(X, \|\cdot\|_X)$ be a normed vector space. If $\mathcal{L} \in B(\ell^1_\nu, X)$, then

$$\|\mathcal{L}\|_{B(\ell^1_\nu, X)} = \sup_{k \in \mathbb{N}_0} \|\mathcal{L}(\xi_k)\|_X.$$  

**Proof.** It is clear that

$$\|\mathcal{L}\|_{B(\ell^1_\nu, X)} \geq \sup_{k \in \mathbb{N}_0} \|\mathcal{L}(\xi_k)\|_X,$$

since $\|\xi_k\|_\nu = 1$ for all $k \in \mathbb{N}_0$ by definition.

Conversely, let $a \in \ell^1_\nu$ be arbitrary and observe that

$$a = a_0 \xi_0 + 2 \sum_{k=1}^{\infty} a_k \varepsilon_k \nu^k.$$  

Therefore, since $\mathcal{L}$ is bounded,

$$\mathcal{L}(a) = a_0 \mathcal{L}(\xi_0) + 2 \sum_{k=1}^{\infty} a_k \mathcal{L}(\xi_k) \nu^k.$$  

Consequently,

$$\|\mathcal{L}(a)\|_X \leq \sup_{k \in \mathbb{N}_0} \|\mathcal{L}(\xi_k)\|_X \|a\|_\nu$$

for any $a \in \ell^1_\nu$, which proves the claim. \qed
Now, suppose $L \in \mathcal{B}(\ell_{\nu_1}^1, \ell_{\nu_2}^1)$, where $\nu_1, \nu_2 > 1$. Then $L$ can be identified with an infinite dimensional matrix, with respect to the basis $(e_k)_{k \in \mathbb{N}_0}$, in the usual way. More precisely, there exists unique coefficients $\{L_{ij} \in \mathbb{C} : i, j \in \mathbb{N}_0\}$ such that

$$L(e_j) = \sum_{i=0}^{\infty} L_{ij} e_i \simeq \begin{bmatrix} L_{00} & L_{01} & \cdots \\ L_{10} & L_{11} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}^T, \quad j \in \mathbb{N}_0.$$  

Hence

$$L(a) = \begin{bmatrix} L_{00} & L_{01} & \cdots \\ L_{10} & L_{11} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \end{bmatrix}, \quad (3.6)$$

for any $a \in \ell_{\nu_1}^1$. In this particular setting, Proposition 3.2.8 can be interpreted as the statement that $\|L\|_{\mathcal{B}(\ell_{\nu_1}^1, \ell_{\nu_2}^1)}$ is a weighted supremum of the $\ell_{\nu_2}^1$-norms of the columns of $L$. Moreover, in this case the converse of Proposition 3.2.8 holds as well:

**Proposition 3.2.9.** Let $\nu_1, \nu_2 > 1$ and suppose $\{L_{ij} \in \mathbb{C} : i, j \in \mathbb{N}_0\}$ are coefficients such that the expression in (3.6) yields a well-defined linear operator $L : \ell_{\nu_1}^1 \to \mathbb{C}^{\mathbb{N}_0}$, i.e., $(L(a))_k$ is finite for all $a \in \ell_{\nu_1}^1$ and $k \in \mathbb{N}_0$. Then $L \in \mathcal{B}(\ell_{\nu_1}^1, \ell_{\nu_2}^1)$ if and only if $\sup_{l \in \mathbb{N}_0} \varepsilon_{l, \nu_1} \|L(\cdot, l)\|_{\nu_2} < \infty$. Moreover, if $L \in \mathcal{B}(\ell_{\nu_1}^1, \ell_{\nu_2}^1)$, then

$$\|L\|_{\mathcal{B}(\ell_{\nu_1}^1, \ell_{\nu_2}^1)} = \sup_{l \in \mathbb{N}_0} \varepsilon_{l, \nu_1} \|L(\cdot, l)\|_{\nu_2}. \quad (3.7)$$

**Proof.** It follows directly from Proposition 3.2.8 that $\sup_{l \in \mathbb{N}_0} \varepsilon_{l, \nu_1} \|L(\cdot, l)\|_{\nu_2} < \infty$ whenever $L \in \mathcal{B}(\ell_{\nu_1}^1, \ell_{\nu_2}^1)$ and that in this case the operator norm is given by (3.7). Conversely, suppose $\sup_{l \in \mathbb{N}_0} \varepsilon_{l, \nu_1} \|L(\cdot, l)\|_{\nu_2} < \infty$. Let $a \in \ell_{\nu_1}^1$ be arbitrary, then

$$\|L(a)\|_{\nu_2} \leq \sum_{l=0}^{\infty} |L_{0l}| |a_l| + 2 \sum_{k=1}^{\infty} \sum_{l=0}^{\infty} |L_{kl}| |a_l| \nu_2^k$$

$$= \sum_{l=0}^{\infty} |a_l| \|L(\cdot, l)\|_{\nu_2}$$

$$= |a_0| \|L(\cdot, 0)\|_{\nu_2} + 2 \sum_{l=1}^{\infty} \varepsilon_{\nu_1, l} |a_l| \|L(\cdot, l)\|_{\nu_2} \nu_1^l$$

$$\leq \left( \sup_{l \in \mathbb{N}_0} \varepsilon_{l, \nu_1} \|L(\cdot, l)\|_{\nu_2} \right) \|a\|_{\nu_1}.$$  

Therefore, since the latter quantity is finite, $L : \ell_{\nu_1}^1 \to \ell_{\nu_2}^1$ is a bounded linear operator. \qed
3.2.3 Multivariate sequences

In this section we introduce a space of sequences indexed by \(d\)-dimensional multi-indices, where \(d \in \mathbb{N}\). This space will be used to analyze Taylor series of analytic functions \(P : \{ z \in \mathbb{C} : |z| \leq \nu \}^d \to \mathbb{C}^n \), where \(\nu > 0\). Such functions arise in the analysis of local charts on (un)stable manifolds via the parameterization method developed in [18].

Formally, a sequence indexed by \(d\)-dimensional multi-indices is a function \(p : \mathbb{N}_0^d \to \mathbb{C}^n\). The function \(p\) is usually referred to as a \(d\)-dimensional array or multivariate sequence. In analogy with ordinary sequences, we shall write (as usual)

\[ p_k := p_{k_1 \ldots k_d} := p(k), \quad k \in \mathbb{N}_0^d. \]

Furthermore, for any multi-index \(k \in \mathbb{N}_0^d\), we write \(|k| = \sum_{i=1}^d k_i\), which is not to be confused with the absolute value of a (complex) number. In addition, we introduce a partial ordering \(\preceq\) on \(\mathbb{N}_0^d\) by

\[ k \preceq l \quad \text{def} \quad \Leftrightarrow \quad k_j \leq l_j, \quad \forall \ 1 \leq j \leq d. \]

We will now follow the same approach as in the previous section to set up a functional analytic framework for analyzing geometrically decaying arrays. Let \(\nu > 0\) and define

\[ W_{1,\nu,n,d} := \left\{ p : \mathbb{N}_0^d \to \mathbb{C}^n \left| \sum_{k \in \mathbb{N}_0^d} \left| p_k \right| |k| \nu^{|k|} < \infty, \ 1 \leq j \leq n \right\} \]

endowed with the norm

\[ \| p \|_{W_{1,\nu,n,d}} := \max_{1 \leq j \leq n} \sum_{k \in \mathbb{N}_0^d} \left| p_k \right| |k| \nu^{|k|}. \]

In the case that the dimension \(d\) can be easily inferred from the context it will be omitted from the notation. In addition, if \(n = 1\) and it is clear from the context whether \(p \in \ell_{\nu}^1\) or \(p \in W_{1,\nu}^1\), we shall write \(\| p \|_{W_{1,\nu,n,d}} = \| p \|_{\nu}\).

Next, recall that the Taylor coefficients of the product of two Taylor series is given by the one-sided discrete convolution, also referred to as the Cauchy product. More precisely, if \(f, g : \{ z \in \mathbb{C} : |z| < \nu \}^d \to \mathbb{C}\) admit power series expansions

\[ f(z) = \sum_{k \in \mathbb{N}_0^d} f_k z^k, \quad g(z) = \sum_{k \in \mathbb{N}_0^d} g_k z^k, \]

where \(\tilde{f} = (f_k)_{k \in \mathbb{N}_0^d}\) and \(\tilde{g} = (g_k)_{k \in \mathbb{N}_0^d}\) are \(d\)-dimensional arrays, then

\[ (fg)(z) = \sum_{k \in \mathbb{N}_0^d} \left( \tilde{f} \ast \tilde{g} \right)_k z^k, \quad \left( \tilde{f} \ast \tilde{g} \right)_k := \sum_{\alpha,\beta \in \mathbb{N}_0^d} \tilde{f}_\alpha \tilde{g}_\beta, \quad (3.8) \]

on \(\{ z \in \mathbb{C} : |z| < \nu \}^d\). In particular, the Cauchy product \(\ast\) yields a natural product structure on \(W_{1,\nu}^1\). This is summarized in the following proposition.
Proposition 3.2.10. The space $(W_1^1, *)$ is a commutative Banach algebra.

Proof. This follows directly from the definition of $*$ in (3.8) and the triangle inequality. □

Next, we derive an expression for the norm of a bounded linear operator on $W_1^1$. For this purpose we introduce a multivariate analog of the corner-points:

Definition 3.2.11. The corner points $\{\xi_{k,d,\nu}\}_{k\in \mathbb{N}_0^d} \subset W_1^1$ of the unit ball in $W_1^1$ are defined by $(\xi_{k,d,\nu})_l := \nu^{-|k|} \delta_{kl}$, where $l \in \mathbb{N}_0^d$. We shall write $\xi_{k,d,\nu} = \xi_k$ whenever there is no chance of confusion.

As before, the norm of a bounded linear operator on $W_1^1$ can be computed by evaluating it at the corner-points.

Proposition 3.2.12. Let $(X, \|\cdot\|_X)$ be a normed vector space. If $L \in B(W_1^1, X)$, then

$$
\|L\|_{B(W_1^1, X)} = \sup_{k\in \mathbb{N}_0^d} \|L(\xi_k)\|_X .
$$

Proof. See Proposition 3.2.8. □

3.3 An equivalent zero finding problem

In this section we set up a zero finding problem for establishing the existence of connecting orbits. Let us start by giving a precise description of the problem. Suppose $\tilde{p}_0, \tilde{q}_0 \in \mathbb{R}^n$ are hyperbolic equilibria of $g$. The objective is to validate an isolated connecting orbit $u$ from $\tilde{p}_0$ to $\tilde{q}_0$, which is robust with respect to “small” perturbations in $g$, by solving a boundary value problem (BVP) on a finite time domain. The method is based on the observation that a connecting orbit from $\tilde{p}_0$ to $\tilde{q}_0$ is characterized by

$$
\begin{cases}
\frac{du}{dt} = g(u), & t \in [0, L], \\
u(0) \in W_{loc}^u(\tilde{p}_0), \\
u(L) \in W_{loc}^s(\tilde{q}_0),
\end{cases}
$$

(3.9)

where $L > 0$ is the time of flight needed to travel from $W_{loc}^u(\tilde{p}_0)$ to $W_{loc}^s(\tilde{q}_0)$.

If $W^u(\tilde{p}_0)$ and $W^s(\tilde{q}_0)$ intersect transversally along $u$, then the connecting orbit is robust, i.e., it will persist for sufficiently “small” perturbations in $g$. In this case, the intersection $W^u(\tilde{p}_0) \cap W^s(\tilde{q}_0) \cap U$, where $U$ is a neighborhood of the connecting orbit in which it is unique, is necessarily an one dimensional manifold. Hence, by counting dimensions, a necessary condition for the existence of a transverse isolated connecting orbit is

$$
n_u + n_s - n = 1, \quad n_u := \dim W^u(\tilde{p}_0), \quad n_s := \dim W^s(\tilde{q}_0) .
$$
This condition is often referred to as a *non-degeneracy* condition for connecting orbits. We shall henceforth assume that this condition is satisfied. In particular, we do not assume a-priori that the connecting orbit is isolated and transverse. Instead, we will obtain these properties from the proof of existence (a contraction argument), see Proposition 3.3.19.

We start by setting up equations for local charts on the (un)stable manifolds by using the parameterization method [18] and the methodology presented in [75]. These charts will be used to supplant the boundary conditions in (3.9) with explicit equations. Next, we set up an equivalent system of equations for the differential equation by using Chebyshev series and domain decomposition as explained in Chapter 2. Finally, in order for the resulting zero finding problem to be well posed, we complete the system of equations by imposing appropriate phase conditions.

### 3.3.1 Charts on the (un)stable manifolds

In this section we give a brief overview of the method developed in [75] to compute local charts on the (un)stable manifolds. The reader is referred to [75] for a more detailed exposition of the theory. To make the discussion more precise, we consider the computation of a local chart on the stable manifold of $\tilde{q}_0$. A chart on the unstable manifold of $\tilde{p}_0$ can be computed in the same way by reversing the sign of the vectorfield.

**The Parameterization Method**  The idea of the parameterization method [18] is to construct a diffeomorphism which conjugates the nonlinear dynamics on the stable manifold to an easier and fully understood flow $\psi$. For the sake of simplicity, let us assume that $Dg(\tilde{q}_0)$ is diagonalizable. This assumption is, however, not necessary, as we will explain in a moment.

Let $\lambda_1^s, \ldots, \lambda_n^s \in \mathbb{C}$ be the stable eigenvalues of $Dg(\tilde{q}_0)$. If all eigenvalues are real and semisimple, then there exists neighborhoods $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^{n_s}$ of $\tilde{q}_0$ and 0, respectively, such that the dynamics on $W^s(\tilde{q}_0) \cap U$ is conjugate to the flow

$$\psi(t, \phi) := \exp(t \cdot \text{diag}(\lambda_1^s, \ldots, \lambda_n^s)) \phi, \quad t \geq 0, \ \phi \in V.$$  

(3.10)

If some of the eigenvalues are complex, however, special care has to be taken. Let us for the moment forget about this technicality and consider the *complex* dynamics generated by $u' = g(u)$ on $\mathbb{C}^n$. Then the dynamics on the complex local stable manifold, which we denote by $W^{s,c}_{\text{loc}}(\tilde{q}_0)$, is conjugate to the flow $\psi$ restricted to the polydisk

$$\mathbb{B}_{\nu_s} := \left\{ \phi \in \mathbb{C}^{n_s} : \max_{1 \leq i \leq n_s} |\phi_i| \leq \nu_s \right\},$$

for some sufficiently small $\nu_s > 0$.

The idea is to find an *analytic* map $Q : \mathbb{B}_{\nu_s} \to \mathbb{C}^n$ which conjugates the nonlinear flow $\varphi$ on $W^{s,c}_{\text{loc}}(\tilde{q}_0)$ to the linear flow $\psi$ on $\mathbb{B}_{\nu_s}$ for $t \geq 0$. In other words, we seek a map $Q$ such that
\[
\begin{array}{ccc}
\mathbb{B}_{\nu_s} & \xrightarrow{Q} & \mathbb{C}^n \\
\Downarrow \psi & & \Downarrow \phi \\
\mathbb{B}_{\nu_s} & \xrightarrow{Q} & \mathbb{C}^n 
\end{array}
\]

commutes, i.e., \(Q(\psi(t, \phi)) = \phi(t, Q(\phi))\) for all \((t, \phi) \in \mathbb{R}_{\geq 0} \times \mathbb{B}_{\nu_s}\). Differentiation of this relation at \(t = 0\) yields the so-called invariance equation:

\[DQ(\phi) \cdot \text{diag}(\lambda^s_1, \ldots, \lambda^s_n) \phi = g(Q(\phi)), \quad \phi \in \mathbb{B}_{\nu_s}. \tag{3.11}\]

Note that this equation does not depend on time or the flow anymore. Moreover, it is easy to see that if the invariance equation holds, then

\[t \mapsto u(t) := Q(\psi(t, \phi))\]

is an orbit in \(W^{s,c}(\tilde{q}_0)\) for any \(\phi \in \mathbb{B}_{\nu_s}\), i.e., \(Q : \mathbb{B}_{\nu_s} \to W^{s,c}_{\text{loc}}(\tilde{q}_0)\) (see \cite[Lemma 2.6]{75}). Therefore, the problem of computing a chart is now reduced to solving (3.11).

**Solving the invariance equation** Since \(Q\) is assumed to be analytic on \(\mathbb{B}_{\nu_s}\), i.e., \(Q\) is analytic on a slightly larger open neighborhood of \(\mathbb{B}_{\nu_s}\), there exist coefficients \(q \in W^{1}_{\nu_s,n,n_s}\) such that

\[Q(\phi) = \sum_{k \in \mathbb{N}_0^d} q_k \phi^k.\]

Observe that the zeroth order Taylor coefficient is necessarily the equilibrium, i.e., \(q_0 = \tilde{q}_0\). Furthermore, since \(Q\) is assumed to be a diffeomorphism, it must hold that

\[DQ(0) : \mathbb{C}^n_s = T_{q_0}W^{s,c}_{\text{loc}}(q_0) = E_s,\]

where \(E_s\) is the stable eigenspace of \(Dg(q_0)\). In fact, the first order Taylor coefficients \(\{q_k : |k| = 1, k \in \mathbb{N}_0^{n_s}\}\) must necessarily be the stable eigenvectors of \(Dg(q_0)\), as we will show in a moment. Note that these are only determined up to a scaling.

To determine the higher order Taylor coefficients \(\{q_k : |k| \geq 2, k \in \mathbb{N}_0^{n_s}\}\), we first introduce the map \(C : W^{1}_{\nu_s,n} \to W^{1}_{\nu_s,n}\) defined by

\[C(w) := \sum_{\alpha \in \mathcal{A}} g_\alpha w^\alpha, \quad w^\alpha := \prod_{j=1}^n [w_j]^{\alpha_j}, \tag{3.12}\]

where the latter product is understood to be the one-sided discrete convolution, and \(\{g_\alpha : \alpha \in \mathcal{A}\} \subset \mathbb{R}^n\), where \(\mathcal{A} \subset \mathbb{N}_0^n\), are the coefficients of \(g\) in the monomial basis. In particular, observe that

\[g(Q(\phi)) = \sum_{k \in \mathbb{N}_0^{n_s}} C_k(q) \phi^k. \tag{3.13}\]
since the Taylor coefficients of the product of two Taylor expansions is given by
the one-sided discrete convolution. Formally, we should incorporate the weight \( \nu_s \)
and dimension \( n_s \) into the notation for \( C \). However, since these parameters can
usually be inferred from the context and we wish to use the same notation for the
unstable manifold, we have chosen to omit it from the notation.

Substitution of the Taylor expansion for \( Q \) into (3.11) yields the follow system
of equations:

\[
\langle \lambda^s, k \rangle q_k - C_k (q) = 0, \quad \lambda^s := \left[ \lambda^s_1 \ldots \lambda^s_{n_s} \right]^T, \quad |k| \geq 2, \tag{3.14}
\]

where \( \langle \cdot, \cdot \rangle \) denotes the standard Hermitian inner product on \( \mathbb{C}^{n_s} \). We shall use
this system of equations to set up a zero finding problem for computing a chart
on \( W_{1,\nu,s}^{\ast,c} (\tilde{q}_0) \). Now, before we proceed, observe that (3.14) is equivalent to

\[
\left[ Dw (q_0) - (\lambda^s, k) I \right] q_k = Dw (q_0) q_k - C_k (q), \quad |k| \geq 1. \tag{3.15}
\]

Moreover, differentiation of (3.13) at \( \phi = 0 \) shows that \( Dw (q_0) q_k = C_k (q) \) for
\( |k| = 1. \) Hence (3.15) reduces to the eigenvalue/eigenvector equation for \( Dw (q_0) \)
for \( |k| = 1. \) Similarly, repeated differentiation of (3.13) at \( \phi = 0 \) shows that the
righthand-side of (3.15) only depends on Taylor coefficients of order strictly below
\( |k| \). In conclusion, the Taylor coefficients \( q \) can be computed recursively up to any
desired order provided \( \langle \lambda^s, k \rangle \neq \lambda_i^s \) for all \( 1 \leq i \leq n_s \) and \( |k| \geq 2. \)

The latter condition is usually referred to as a non-resonance condition and
is related to the regularity of \( Q \). More precisely, in the presence of a resonance,
the parameterization method, as applied above with the linear “model” flow \( \psi \),
does not yield an analytic conjugation \( Q \) anymore. It is explained in \[75\] how to
construct an analytic conjugation in the present of a resonance. The idea is to use a nonlinear
normal form for \( \psi \) instead of just the linear flow in (3.10). The
interested reader is referred to \[75\] for a detailed exposition of the resonant case.
For the sake of presentation, however, we shall assume throughout this paper that
there are no resonances.

We are now ready to set up a zero finding problem for computing the Taylor co-
efficients of \( Q \) (which include the equilibrium and eigenvectors) and the eigenvalues \( \lambda^s \):

**Definition 3.3.1** (Taylor map for stable manifolds). Let \( 0 < \tilde{\nu}_s < \nu_s \) be given
weights. The Taylor map \( F_Q : \mathbb{C}^{n_s} \times W_{1,\nu,s,n}^1 \to W_{1,\tilde{\nu}_s,n}^1 \) for stable manifolds is defined by

\[
(F_Q (\lambda^s, q))_k :=
\begin{cases}
g(q_0), & k = 0, \\
[ Dw (q_0) - (\lambda^s, k) I ] q_k, & |k| = 1, \\
(\lambda^s, k) q_k - C_k (q), & |k| \geq 2.
\end{cases}
\]

**Remark 3.3.2.** The latter map is well-defined since \( (\lambda^s, k) q_k \in W_{1,\tilde{\nu}_s,n}^1 \) for
any \( q \in W_{1,\nu,s,n}^1 \) and \( 0 < \tilde{\nu}_s < \nu_s \).
Remark 3.3.3. If \((\lambda^s, q)\) is a zero of \(F_{Q}\), then so is \((\lambda^s, \mu q)\), where \(\mu \in \mathbb{C}^{n_s}\) and \((\mu q)_k := \mu^k q_k\), see [75, Lemma 2.2]. We will get rid of this extra degree of freedom by fixing the orientation and length of the eigenvectors. In particular, observe that the scaling of the eigenvectors (and in turn the “scaling” of \(q\)) determines the decay rate of the coefficients \(q\) and hence the size of \(B_{\nu}^{u}\). In effect, the length of the eigenvectors determine (roughly speaking) the “size” of the patch on \(W_{loc}^{s,c} (q_0)\) parameterized by \(Q\). The interested reader is referred to [14] for a more thorough explanation where this phenomena is explored in detail.

The zero finding problem for the computation of a chart on the unstable manifold is set up in an analogous way. For the sake of completeness (and introducing notation) let us explicitly state the assumptions and the associated zero finding map. We assume that \(Dg (\tilde{p}_0)\) is diagonalizable and that the associated eigenvalues \(\lambda^u \in \mathbb{C}^{n_u}\) satisfy the non-resonance condition. The goal is to compute a parameterization \(P : \mathbb{B}_{\nu_u} \subset \mathbb{C}^{n_u} \to W_{loc}^{u,c} (\tilde{p}_0)\) of the form \(P(\theta) = \sum_{k \in \mathbb{N}^{n_u}} p_k \theta^k\), where \(\nu_u > 0\), by finding a zero of the following map:

**Definition 3.3.4** (Taylor map for unstable manifolds). Let \(0 < \tilde{\nu}_u < \nu_u\) be given weights. The Taylor map \(F_P : \mathbb{C}^{n_u} \times W_{\nu_u,n}^1 \to W_{\nu_u,n}^1\) for unstable manifolds is defined by

\[
(F_P (\lambda^u, p))_k := \begin{cases} 
g (p_0), & k = 0, \\
[Dg (p_0) - \langle \lambda^u, k \rangle I] p_k, & |k| = 1, \\
\langle \lambda^u, k \rangle p_k - C_k (p), & |k| \geq 2.
\end{cases}
\]

**Symmetry** In the preceding exposition we considered the complex dynamical system \(u' = g(u)\) on \(\mathbb{C}^n\). Our main interest, however, is the computation of invariant manifolds in the real-valued dynamical system on \(\mathbb{R}^n\). We will now explain how we can recover charts for the (un)stable manifolds in the real system from the complex ones through the use of symmetry. We remark that one could also have set up the parameterization method in the real-valued setting from the start. However, in that case, we would have had to separate the cases between the presence of complex eigenvalues and a completely real spectrum. It is in our opinion more convenient from both a practical and theoretical point of view to develop a unified approach.

Let us consider the stable manifold again. Observe that complex eigenvalues will always appear in conjugate pairs, since \(g\) is a polynomial with real coefficients. Suppose there are \(d_s\) complex conjugate pairs of eigenvalues and \(n_s - 2d_s\) real ones. Furthermore, assume that we have ordered the eigenvalues \(\lambda^s\) in such a way that \(\lambda_i^s = \overline{\lambda_{i+1}^s}\) for \(i \in \{2l + 1 : 0 \leq l \leq d_s - 1\}\). Next, define the map \(\Sigma : \mathbb{C}^{n_s} \to \mathbb{C}^{n_s}\) by

\[
\Sigma (z_1, \ldots, z_{2d_s}, z_{2d_s+1}, \ldots, z_{n_s}) := (\overline{z_2}, \overline{z_1}, \ldots, \overline{z_{2d_s}}, \overline{z_{2d_s-1}}, z_{2d_s+1}, \ldots, z_{n_s}),
\]
and note that $\Sigma$ is an involution on $\mathbb{C}^{n_s}$. For this reason we shall frequently write $z^* := \Sigma(z)$. In particular, note that we ordered the stable eigenvalues in such a way that $(\lambda^s)^* = \lambda^s$. Finally, we extend this notion of involution to $W^1_{\nu_s, n}$ by defining
\[(q^*)_k := \overline{q}_k, \quad k \in \mathbb{N}^{n_s}_0.\]

The key observation for obtaining charts for the real manifolds is stated in the following lemma. The proof can be found in [75, Lemma 2.1].

**Proposition 3.3.5.** If $q \in W^1_{\nu_s, n}$ is symmetric, i.e., $q^* = q$, then the map $Q : \mathbb{B}_{\nu_s} \to \mathbb{C}^n$ defined by
\[Q(\phi) := \sum_{k \in \mathbb{N}^{n_s}_0} q_k \phi^k\]
is real valued on the set $\mathbb{B}^{\text{sym}}_{\nu_s} := \{ \phi \in \mathbb{B}_{\nu_s} : \phi^* = \phi \}$. In addition, if $\lambda^s \in \mathbb{C}^{n_s}$ is symmetric and $F(\lambda^s, q) = 0$, then $Q|_{\mathbb{B}^{\text{sym}}_{\nu_s}}$ is a parameterization of the real stable manifold $W^s_{\nu_s}(q_0)$.

**Remark 3.3.6.** Note that $\mathbb{B}^{\text{sym}}_{\nu_s}$ is a real manifold of dimension $n_s$. More precisely, we can identify $\mathbb{B}^{\text{sym}}_{\nu_s}$ with the (real) manifold
\[
\mathbb{B}^{\text{sym, re}}_{\nu_s} := \left\{ \phi \in \mathbb{R}^{n_s} : |\phi_{2j-1}|^2 + |\phi_{2j}|^2 \leq \nu_s, \ 1 \leq j \leq d_s, \ |\phi_j| \leq \nu_s, \ 2d_s + 1 \leq j \leq n_s \right\}
\]
by using the (linear) map $\iota_{\nu_s} : \mathbb{R}^{n_s} \to \mathbb{C}^{n_s}$ defined by
\[
\iota_{\nu_s}(\phi) := \left( \phi_1 + i\phi_2, \phi_1 - i\phi_2, \ldots, \phi_{2d_s-1} + i\phi_{2d_s}, \phi_{2d_s-1} - i\phi_{2d_s}, \phi_{2d_s+1}, \ldots, \phi_{n_s} \right).
\]

**Remark 3.3.7.** Strictly speaking, the assumption that $\lambda^s$ is symmetric is not necessary. To see this, let $e^s_i \in \mathbb{C}^{n_s}$ be the unit vector defined by $[e^s_i]_j = \delta_{ij}$, where $1 \leq i, j \leq n_s$. If $F(\lambda^s, q) = 0$ for some $\lambda^s \in \mathbb{C}^{n_s}$, then
\[Dg(q_0) q_k = \lambda^s_k q_k, \quad |k| = 1.
\]

In particular, if we take the complex conjugate of the left-hand-side of the above expression for $k = e^s_{2j-1}$, we obtain
\[
\overline{Dg(q_0) q_{e^s_{2j-1}}} = Dg(q_0) \overline{q_{e^s_{2j-1}}} = Dg(q_0) q_{e^s_{2j}} = \lambda^s_{2j} q_{e^s_{2j}}, \quad 1 \leq j \leq d_s,
\]
since $g$ is a polynomial with real coefficients and $q^* = q$ by assumption. On the other hand,
\[
\overline{Dg(q_0) q_{e^s_{2j-1}}} = \overline{\lambda^s_{2j-1} q_{e^s_{2j-1}}} = \lambda^s_{2j-1} q_{e^s_{2j}}, \quad 1 \leq j \leq d_s.
\]
Hence it follows that $(\lambda^s)^* = \lambda^s$. 

86
In conclusion, in order to conclude that a point lies on the real stable manifold, it suffices to verify that \( q^* = q \). It is explained in Section 3.4 how we can verify this in practice. For now, let us mention that the verification is based on the following observation whose proof can be found in [75, Lemma 4.1]:

**Lemma 3.3.8.** The map \( F_Q \) is compatible with \( * \), i.e., \( F_Q ((\lambda^*)^*, q^*) = F_Q (\lambda^*, q)^* \) for any \( (\lambda^*, q) \in \mathbb{C}^{n_\lambda} \times W^{1}_{\nu, n} \).

Analogous results hold for the parameterization of the unstable manifold of \( \tilde{p}_0 \).

To avoid clutter in the notation we shall denote the involution associated to the unstable manifold of \( \tilde{p}_0 \) by \( * \) as well.

### 3.3.2 Chebyshev series and domain decomposition

In this section we use the methodology developed in Chapter 2 to recast the differential equation into an equivalent zero finding problem on \( \ell^1_{\nu, n} \). The reader is referred to Chapter 2 for the details. Let \( P_m := \{ t_0 = 0 < t_1 < \ldots < t_m = 1 \} \), where \( m \in \mathbb{N} \), be any partition of \([0, 1]\). Then the differential equation in (3.9) is equivalent to

\[
(P_1) \begin{cases} 
\frac{du_1}{dt} = Lg(u_1), & t \in [0, t_1], \\
\frac{du_i}{dt} = Lg(u_i), & t \in [t_{i-1}, t_i], \\
u_i(t_{i-1}) = u_{i-1}(t_{i-1}), & 2 \leq i \leq m
\end{cases}
\]

where \( 2 \leq i \leq m \). If \( (P_i)_{i=1}^m \) admits a solution, then each \( u_i \) is real-analytic since \( g \) is. Therefore, there exists weights \( \nu_i > 1 \) and real coefficients \( a^i \in \ell^1_{\nu_i, n} \) such that

\[
u_i = a^i_0 + 2 \sum_{k=1}^\infty a^i_k T^i_k
\]

in \( C ([t_{i-1}, t_i]) \). Here \( (T^i_k)_{k \in \mathbb{N}_0} \) are the *shifted* Chebyshev-polynomials on \([t_{i-1}, t_i]\) defined by

\[
T^i_k(t) = T_k \left( \frac{2t - t_{i-1} - t_i}{t_i - t_{i-1}} \right), \quad k \in \mathbb{N}_0.
\]

Next, define the map \( c : \ell^1_{\nu_i, n} \rightarrow \ell^1_{\nu_i, n} \) by

\[
c(a) := \sum_{\alpha \in A} g_\alpha a^\alpha, \quad a^\alpha := \prod_{j=1}^n [a]_{j}^{\alpha_j}, \quad (3.16)
\]

where the latter product is understood to be the symmetric discrete convolution \( * \). Then

\[
g(u_i) = c_0 (a^i) + 2 \sum_{k=1}^\infty c_k (a^i) T^i_k,
\]

87
since the Chebyshev coefficients of the product of two functions is given by the symmetric discrete convolution. Formally, we should incorporate the index $i$ into the notation for $c$ to emphasize its dependence on the weight $\nu_i$. However, since the domain of $c$ can usually be easily inferred from the context, we have chosen to omit the index from the notation.

**Remark 3.3.9.** Throughout this paper we will need to analyze $Dc(a^i)$, where $a^i \in \ell_{\nu_i,n}^1$ and $1 \leq i \leq m$, on numerous occasions. For this reason, we state here for future reference how this derivative can be computed in an efficient way. Since $(\ell_{\nu_i,n}, *)$ is a Banach algebra, we may use the “usual” rules of calculus to compute the derivative of $c$. In particular, direct differentiation of (4.19) with respect to $a^i$ shows that

$$D [c]_j (a^i) \bar{a}^i = \sum_{l=1}^{n} \hat{g}^{ijl} \ast [\bar{a}]_l, \quad \bar{a}^i \in \ell_{\nu_i,n}^1, \quad 1 \leq j \leq n,$$

(3.17)

where $[c]_j$ denotes the $j$-th component of $c$ and $\hat{g}^{ijl} \in \ell_{\nu_i}^1$ are the Chebyshev coefficients of

$$\frac{\partial g_{ij}}{\partial x_l} \left( a^i_0 + 2 \sum_{k=1}^{\infty} a^i_k T_k^l \right), \quad 1 \leq l \leq n.$$  

(3.18)

Substitution of the Chebyshev expansions for $(u_i)_{i=1}^m$ into $(P_i)_{i=1}^m$ yields an equivalent system of equations for the coefficients and gives rise to the following map:

**Definition 3.3.10** (Chebyshev map for ODEs). Let $(\nu_i)_{i=1}^m$ and $(\bar{\nu}_i)_{i=1}^m$ be collections of weights such that $1 < \bar{\nu}_i < \nu_i$ for all $1 \leq i \leq m$. The Chebyshev map for ODEs is the function $F_u : \bigoplus_{i=1}^m \ell_{\nu_i,n}^1 \to \ell_{\bar{\nu}_i,n}^1 / \mathbb{C}^n \oplus \bigoplus_{i=2}^m \ell_{\nu_i,n}^1$ defined by

$$F_u (a) := (f_1 (a^1) , f_2 (a^2) , \ldots , f_m (a^{m-1}, a^m)),$$

where $a = (a^1, \ldots, a^m)$, $f_1 : \ell_{\nu_1,n}^1 \to \ell_{\bar{\nu}_1,n}^1 / \mathbb{C}^n$ is given by

$$(f_1 (a^1))_k := ka^1_k - \frac{L (t_1 - t_0)}{4} (c_{k-1} (a^1) - c_{k+1} (a^1)),$$

$k \in \mathbb{N}$,

and $f_i : \ell_{\nu_{i-1},n}^1 \times \ell_{\nu_i,n}^1 \to \ell_{\nu_i,n}^1$ by

$$f_i (a^{i-1}, a^i) :=
\begin{cases}
  a^i_0 - a^i_0 + 2 \sum_{l=1}^{\infty} ((-1)^l a^i_l - a^i_l), & k = 0, \\
  ka^i_k - \frac{L (t_i - t_{i-1})}{4} (c_{k-1} (a^i) - c_{k+1} (a^i)), & k \in \mathbb{N},
\end{cases}$$

for $2 \leq i \leq m$.

**Remark 3.3.11.** The map $F_u$ is well-defined, since $(ka^i_k)_{k \in \mathbb{N}_0} \in \ell_{\bar{\nu}_i,n}^1$ for any $a^i \in \ell_{\nu_i,n}^1$ and $1 < \bar{\nu}_i < \nu_i$.  

88
Let us stress the subtle difference between the latter map and the one constructed in Chapter 2 in the current setting we allow for complex Chebyshev coefficients. The main reason for this is that the parameterization maps $P$ and $Q$ (see the previous section) are in principle complex-valued and we wish to use them to prove that $u(0) \in W^u_{\text{loc}}(\tilde{p}_0)$ and $u(L) \in W^s_{\text{loc}}(\tilde{q}_0)$. We will conclude a-posteriori that the Chebyshev coefficients are in fact real by invoking symmetry arguments. To be more precise, for any $a = (a^1, \ldots, a^m)$, define $\overline{a} := \left(\overline{a^1}, \ldots, \overline{a^m}\right)$ by $(\overline{a^j})_k := \overline{a^j_k}$. We will conclude that an element $a$ is real, i.e., $\overline{a} = a$, by using the following observation:

**Lemma 3.3.12.** The map $F_u$ is compatible with conjugation, i.e., $F_u(\overline{a}) = \overline{F_u(a)}$.

**Proof.** Let $a = (a^1, \ldots, a^m) \in \prod_{j=1}^m \ell^1_{\nu_j, n}$ be arbitrary. We will prove that $c(\overline{a^j}) = \overline{c(a^j)}$ for $1 \leq j \leq m$. The desired result follows directly from this observation. Recall that a Chebyshev series is a Fourier series up to coordinate transformation. To be more precise, let $1 \leq j \leq m$, set $a^j_k := a^j_k$ for $k \in \mathbb{N}$, and define

$$s_j(\theta) := \frac{t_j - t_{j-1}}{2} (\cos \theta + 1) + t_{j-1}, \quad \theta \in [0, \pi].$$

Then

$$a^j_0 + 2 \sum_{k=1}^{\infty} a^j_k T_k^j(s_j(\theta)) = \sum_{k \in \mathbb{Z}} a^j_k e^{ik\theta}, \quad \theta \in [0, \pi],$$

by definition of the Chebyshev polynomials (see Definition 3.2.1). Note that the latter series converges uniformly to an analytic $2\pi$-periodic function, since $a^j \in \ell^1_{\nu_j, n}$. Furthermore, since $a^j_{-k} = a^j_k$ (by definition), it follows that

$$g\left(a^j_0 + 2 \sum_{k=1}^{\infty} a^j_k T_k^j\right) = \sum_{k \in \mathbb{Z}} c_k(a^j) e^{ik\theta},$$

where $c$ is defined in (4.19) and we have set $c_{-k}(a^j) := c_k(a^j)$ for $k \in \mathbb{N}$. In particular,

$$g\left(a^j_0 + 2 \sum_{k=1}^{\infty} a^j_k T_k^j\right) = \sum_{k \in \mathbb{Z}} c_k(a^j) e^{ik\theta}.$$

On the other hand, since $g$ is real-analytic, the same reasoning shows that

$$g\left(\overline{a^j_0 + 2 \sum_{k=1}^{\infty} a^j_k T_k^j}\right) = g\left(\overline{a^j_0 + 2 \sum_{k=1}^{\infty} a^j_k T_k^j}\right) = \sum_{k \in \mathbb{Z}} c_k(a^j) e^{ik\theta}.$$

Therefore, since a (pointwise) convergent Fourier series is unique, we conclude that $c(\overline{a^j}) = \overline{c(a^j)}$. \qed

89
Finally, observe that by construction we now have the following result:

**Proposition 3.3.13.** Suppose \( a \in \bigoplus_{i=1}^m \ell_{\nu}^1 \) is symmetric, i.e., \( \overline{a} = a \), then \( F(a) = 0 \) if and only if the functions \( \{ u_i = a_i^0 + 2 \sum_{k=1}^{\infty} a_k^i T_k^i : 1 \leq i \leq m \} \) constitute a solution of \((P_i)_{i=1}^m\).

### 3.3.3 The connecting orbit map

In this section we set up a zero finding problem for (3.9). We have already set up appropriate zero finding mappings for the ODE and charts on the (un)stable manifolds. What remains are appropriate phase space conditions.

**Boundary conditions** We can now replace the boundary conditions in (3.9) with explicit equations. Let \( P \) and \( Q \) denote the local parameterizations of the complex unstable and stable manifold as before, respectively. Then the conditions \( u(0) \in W^u_{\text{loc}}(\tilde{p}_0) \) and \( u(1) \in W^s_{\text{loc}}(\tilde{q}_0) \) are equivalent to the problem of finding coordinates \( \theta \in \mathbb{B}^\text{sym}_{\nu_u} \) and \( \phi \in \mathbb{B}^\text{sym}_{\nu_s} \) such that

\[
\begin{align*}
\left. u_1(0) - P(\theta) = a_0^1 + 2 \sum_{k=1}^{\infty} (-1)^k a_k^1 \theta_k = 0, \right.
\end{align*}
\]

\[
\begin{align*}
\left. u_m(1) - Q(\phi) = a_0^m + 2 \sum_{k=1}^{\infty} a_k^m \phi_k = 0. \right.
\end{align*}
\]

As mentioned before, we will verify a-posteriori that \( \theta^* = \theta \) and \( \phi^* = \phi \) so that \( P(\theta) \) and \( Q(\theta) \) are points on the real (un)stable manifolds.

**Length of the eigenvectors** Recall that the first order Taylor coefficients of \( P \) and \( Q \) are determined up to a rescaling. To get rid of this extra degree of freedom we prescribe the length and orientation of the eigenvectors of \( Dg(\tilde{p}_0) \) and \( Dg(\tilde{q}_0) \). More precisely, we require that

\[
\begin{align*}
\langle p_k, \hat{p}_k \rangle - \epsilon_{u,k} &= 0, \quad |k| = 1, \ k \in \mathbb{N}^n_u, \\
\langle q_k, \hat{q}_k \rangle - \epsilon_{s,k} &= 0, \quad |k| = 1, \ k \in \mathbb{N}^n_s,
\end{align*}
\]

where \( \hat{p}_k, \hat{q}_k \in \mathbb{C}^n \) are prescribed vectors and \( \epsilon_{u,k}, \epsilon_{s,k} > 0 \). In practice, \( \hat{p}_k \) and \( \hat{q}_k \) are numerical approximations of the eigenvectors of \( Dg(\tilde{p}_0) \) and \( Dg(\tilde{q}_0) \), respectively, and \( \epsilon_{u,k}, \epsilon_{s,k} \) are their respective squared lengths. In order to respect the symmetry \( \star \), we impose the following ordering:

\[
\begin{align*}
\hat{p}_k^\star &= \overline{p}_k, \quad \epsilon_{u,k^\star} = \epsilon_{u,k}, \quad |k| = 1, \ k \in \mathbb{N}^n_u, \\
\hat{q}_k^\star &= \overline{q}_k, \quad \epsilon_{s,k^\star} = \epsilon_{s,k}, \quad |k| = 1, \ k \in \mathbb{N}^n_s.
\end{align*}
\]

We recall again that the length of the eigenvectors determines the decay rate of the Taylor coefficients and hence the size of the domains of \( P \) and \( Q \).
Translation invariance in time  Finally, we introduce a phase condition to fix
the time parameterization of the connecting orbit. In [12,29] a phase condition
specifically tailored for continuation is presented. The idea is to fix a reference
function \( \tilde{u} \) and to minimize the functional

\[
s \mapsto \int_{-\infty}^{\infty} \| u(t - s) - \tilde{u}(t) \|^2 \, dt
\]
on some appropriate functions space, where \( u \) is the connecting orbit. This phase
condition is used in popular software packages for continuation such as AUTO
and MATCONT and was first suggested in [29].

The intuition is that this phase condition enforces the connecting orbit to re-
main as close as possible (in the \( L^2 \)-sense) to the reference solution with respect to
small shifts in time. In practice, \( \tilde{u} \) is the solution computed at the previous contin-
uation step (or just the numerical approximation \( \hat{u} \) in case we are not performing
continuation). In particular, a necessary condition for the latter functional to have
a minimum at \( s = 0 \) is

\[
\int_{-\infty}^{\infty} \langle u(t) - \tilde{u}(t), \tilde{u}'(t) \rangle \, dt = 0. \tag{3.21}
\]

We shall use \((3.21)\) to construct an appropriate phase condition in terms of
Chebyshev coefficients by approximating the integral on a finite domain. First,
write

\[
u|_{[-1,1]} = \sum_{i=1}^{m} u_i 1_{[t_{i-1}, t_i]}|_{[-1,1]}, \quad \tilde{u}|_{[-1,1]} = \sum_{i=1}^{m} \tilde{u}_i 1_{[t_{i-1}, t_i]}|_{[-1,1]},
\]

which correspond to the pieces of \( u \) and \( \tilde{u} \) computed in practice. If the time of
flight \( L > 0 \) is sufficiently large, then

\[
\int_{-\infty}^{\infty} \langle u(t) - \tilde{u}(t), \tilde{u}'(t) \rangle \, dt \approx \int_{-1}^{1} \langle u(t) - \tilde{u}(t), \tilde{u}'(t) \rangle \, dt
\]

\[
= \sum_{i=1}^{m} \int_{t_{i-1}}^{t_i} \langle u_i(t) - \tilde{u}_i(t), \tilde{u}_i'(t) \rangle \, dt.
\]

Next, write

\[
u_i = a_i^0 + 2 \sum_{k=1}^{\infty} a_k^i T_k^i, \quad \tilde{u}_i = b_i^0 + 2 \sum_{k=1}^{\infty} b_k^i T_k^i, \quad 1 \leq i \leq m.
\]

For notational convenience, let us omit the superscripts from the Chebyshev coef-
ficients and assume (for the moment) that \( u_i \) and \( \tilde{u}_i \) are scalar functions. Then

\[
\langle u_i - \tilde{u}_i, \tilde{u}_i' \rangle_{L^2}
\]

\[
= 2 (a_0 - b_0) \sum_{l=1}^{\infty} b_l \left( T_0^i, \frac{dT_l^i}{dt} \right)_{L^2} + 4 \sum_{k,l=1}^{\infty} (a_k - b_k) b_l \left( T_k^i, \frac{dT_l^i}{dt} \right)_{L^2}, \tag{3.22}
\]
where $\langle \cdot, \cdot \rangle_{L^2}$ denotes the standard complex inner product on $L^2([t_{i-1}, t_i])$.

Now, rescale time back to $[-1,1]$, use the coordinate transformation $\theta = \arccos(t)$ and the definition of the Chebyshev polynomials to see that

$$
\langle T^i_k, dT^i_l \rangle_{L^2} = l \int_0^\pi \sin(l \theta) \cos(k \theta) \, d\theta
$$

$$
= \begin{cases} 
\frac{2l^2}{l^2 - k^2}, & k + l \equiv 1 \mod 2, \\
0, & \text{otherwise},
\end{cases}
$$

for any $k, l \in \mathbb{N}_0$. Finally, substitution of the latter expression into (3.22) yields

$$
\langle u_i - \bar{u}_i, \bar{u}_i' \rangle_{L^2} = 4 \left( (a_0 - b_0) \sum_{l=0}^\infty b_{2l+1} + 2 \sum_{s=1}^\infty \sum_{k+l=2s+1} (a_k - b_k) \frac{l^2}{l^2 - k^2} \right).
$$

(3.23)

If $u_i$ and $\bar{u}_i$ are vector-valued, then we need to carry out the above computations component-wise and sum over the components.

In practice, we choose the Chebyshev coefficients of $\bar{u}$ to be real, since in the end we wish to establish the existence of a real-valued connecting orbit. Altogether, this motivates the following definition:

**Definition 3.3.14** (Phase condition for translation invariance in time). Let $b = (b^1, \ldots, b^m) \in \bigoplus_{i=1}^m \ell^{1}_{\nu_i,n}$ be given symmetric sequences, i.e., $\bar{b} = b$, such that $b^i_k = 0$ for $k \geq N_i$, for some $N \in \mathbb{N}^m$. The phase condition for translation invariance in time is the map $\eta : \bigoplus_{i=1}^m \ell^{1}_{\nu_i,n} \to \mathbb{C}$ defined by the following truncated version of (3.23):

$$
\eta (a^1, \ldots, a^m) :=
$$

$$
\sum_{i=1}^m \sum_{j=1}^n \left( [a^i_0 - b^i_0]_j + \sum_{k=0}^{\lfloor \frac{N_i-1}{2} \rfloor - 1} [b^i_{2k+1}]_j + 2 \sum_{s=1}^{N_i-2} \sum_{1 \leq k,l \leq N_i-1} [a^i_k - b^i_k]_j [b^i_l]_j \frac{l^2}{l^2 - k^2} \right).
$$

(3.24)

**Remark 3.3.15.** The expression for $\eta$ might seem complicated at first sight. Note, however, that $\eta$ is really just a linear map.

We are now ready to set up the connecting orbit map. To this end, let

$$
\nu := (\nu_u, \nu_s, \nu_1, \ldots, \nu_m), \quad \bar{\nu} := (\bar{\nu}_u, \bar{\nu}_s, \bar{\nu}_1, \ldots, \bar{\nu}_m)
$$
be given weights such that \( \tilde{\nu} < \nu, \nu_u, \nu_s > 0, \nu_i > 1 \) for \( 1 \leq i \leq m \), and set

\[
\mathcal{X}_\nu := \mathbb{B}_{\nu_u} \times \mathbb{B}_{\nu_s} \times \mathbb{C}^{n_u} \times \mathbb{C}^{n_s} \times \prod_{i=1}^{m} \ell_{\nu_i,n}^{1} \times W_{\nu_u,n,n_u}^{1} \times W_{\nu_s,n,n_s}^{1},
\]

\[
\mathcal{Y}_\tilde{\nu} := \mathbb{C}^{2n+1+n_u+n_s} \times \ell_{\tilde{\nu},n}^{1} / \mathbb{C}^{n} \times \prod_{i=1}^{m} \ell_{\nu_i,n}^{1} \times W_{\nu_u,n,n_u}^{1} \times W_{\nu_s,n,n_s}^{1}.
\]

**Definition 3.3.16** (Chebyshev-Taylor map for connecting orbits). The Chebyshev-Taylor map \( F : \mathcal{X}_\nu \to \mathcal{Y}_\tilde{\nu} \) for connecting orbits is defined by

\[
F(x) := \begin{bmatrix}
  a_0^1 + 2 \sum_{k=1}^{\infty} (-1)^k a_k^1 - \sum_{k \in \mathbb{N}^1_{nu}} p_k \theta^k \\
  a_0^m + 2 \sum_{k=1}^{\infty} a_k^m - \sum_{k \in \mathbb{N}^1_{ns}} q_k \phi^k \\
  \langle p e_k, \hat{p} e_k \rangle - \epsilon_{u,k}^{n_u} \\
  \langle q e_k, \hat{q} e_k \rangle - \epsilon_{s,k}^{n_s} \\
  \eta (a^1, \ldots, a^m) \\
  F_u (a^1, \ldots, a^m) \\
  F_P (\lambda^u, p) \\
  F_Q (\lambda^s, q)
\end{bmatrix},
\]

where \( x = (\theta, \phi, \lambda^u, \lambda^s, a^1, \ldots, a^m, p, q) \).

**Remark 3.3.17.** We shall frequently denote elements in \( \mathcal{Y}_\tilde{\nu} \) by

\[
y = (y_{t_0}, y_{t_m}, y_{\tilde{p}_1}, y_{\tilde{q}_1}, y_\eta, y_a, y_p, y_q),
\]

where

- \( y_{t_0}, y_{t_m} \in \mathbb{C}^n \) correspond to the equations for the boundary conditions at \( t = t_0 \) and \( t = t_m \), respectively,
- \( y_{\tilde{p}_1} \in \mathbb{C}^{n_u}, y_{\tilde{q}_1} \in \mathbb{C}^{n_s} \) correspond to the equations for fixing the length and orientation of the eigenvectors,
- \( y_\eta \in \mathbb{C} \) corresponds to the phase condition for fixing the time parameterization of the orbit,
- \( y_a, y_p \) and \( y_q \) correspond to the Chebyshev and Taylor coefficients, respectively, as before.
The only reason for introducing the weights $\tilde{\nu}$ is to specify the codomain of $F$. These weights are irrelevant though, since we will establish the existence of a connecting orbit by analyzing a fixed point map from $X_{\nu}$ into itself, see Section 3.4.2. For this reason, we only specify a norm on $X_{\nu}$. Namely, we set

$$
\|x\|_{X_{\nu}} := \max \left\{ \max_{1 \leq i \leq n_u} |\theta_i|, \max_{1 \leq i \leq n_s} |\phi_i|, \max_{1 \leq i \leq n_u} |\lambda_i^u|, \max_{1 \leq i \leq n_s} |\lambda_i^s|, \max_{1 \leq i \leq m} \|a_i\|_{\nu_i,n}, \|p\|_{\nu_u,n}, \|q\|_{\nu_s,n} \right\},
$$

where $x = (\theta, \phi, \lambda^u, a^1, \ldots, a^m, p, q)$.

**Symmetry revisited** Next, we examine the compatibility of $F$ with respect to the symmetries introduced in the previous sections. In particular, the involution operations on the space of Taylor and Chebyshev coefficients yield an symmetry operation $\star$ on $X_{\nu}$ defined by

$$
x^{\star} := (\theta^*, \phi^*, (\lambda^u)^*, (\lambda^s)^*, a^*, p^*, q^*).
$$

Similarly, we define an involution on the range $Y_{\tilde{\nu}}$, also denoted by $\star$, via

$$
y^{\star} := (\overline{y_{\nu_{0}}} , \overline{y_{\nu_{m}}} , \overline{y_{\hat{p}_1}} , \overline{y_{\hat{q}_1}} , \overline{y_{\eta}} , \overline{y_{a}} , \overline{y_{\hat{P}}} , \overline{y_{\hat{Q}}} )
$$

**Lemma 3.3.18.** The map $F$ is compatible with $\star$, i.e., $F(x^{\star}) = F(x)^*$ for any $x \in X_{\nu}$.

**Proof.** We start by considering the phase conditions associated to the unstable manifold. First, observe that

$$
a_0^1 + 2 \sum_{k=1}^{\infty} (-1)^k a_k^1 \sum_{k \in N^u_{0}} p_k \theta^k = \overline{a_0^1} + 2 \sum_{k=1}^{\infty} (-1)^k a_k^1 \sum_{k \in N^u_{0}} p_k (\theta^*)^k
$$

by definition of $\star$ and reordering of the series associated to the unstable manifold. Furthermore, note that

$$
\langle p_k, \hat{p}_k \rangle = \overline{\langle p_k, p_k^* \rangle}, \quad |k| = 1, \quad k \in N^u_0,
$$

since $\hat{p}_k$ was ordered in a symmetric way, see (3.19). The computations for the stable manifold are analogous. Next, observe that $\eta(\overline{a}) = \overline{\eta(a)}$, since the Chebyshev coefficients $b$ of the reference orbit $\tilde{u}$ are real. Finally, recall that $F_{\nu}(\overline{a}) = \overline{F_{\nu}(a)}$, $F_{\nu}(\lambda^u)^*, p^*) = F_{\nu}(\lambda^u, p^*)$ and $F_{\nu}(\overline{\lambda^s})^*, q^*) = \overline{F_{\nu}(\lambda^s, q^*)}$ by Lemmas 3.3.12 and 3.3.8, respectively. Altogether, this proves the result. \qed

We are now ready to formulate an appropriate characterization of a connecting orbit:
Proposition 3.3.19. Suppose $F$ has a unique zero $x$ in some open neighborhood $U \subset X$, and assume that $x^* \in U$. Then $p_0, q_0 \in \mathbb{R}^n$ are equilibria of $g$, $P(\theta) \in W^u_{loc}(p_0)$, $Q(\phi) \in W^s_{loc}(q_0)$, and the map $u$ defined by the Chebyshev coefficients is part of an isolated connecting orbit from $p_0$ to $q_0$.

Proof. Suppose $F(x) = 0$, then the previous lemma implies that $x^* = x$, since $x$ is the only zero in $U$ and $x^* \in U$. Consequently, the Chebyshev coefficients $a$ are real and $P(\theta)$ and $Q(\phi)$ are points on the real (un)stable manifolds by Proposition 3.3.5. Therefore, $u$ is part of a connecting orbit from $p_0$ to $q_0$ by Proposition 3.3.13.

Remark 3.3.20. In practice, we seek a zero of $F$ in a closed ball $B_r(\hat{x})$ of radius $r > 0$ centered at an approximate zero $\hat{x}$ obtained through numerical simulation. The numerical computations yield an approximate zero which is almost symmetric (up to machine precision). We manually enforce that $(\hat{x})^* = \hat{x}$ by going through “all” the elements of $\hat{x}$ and imposing the exact symmetry conditions. For example, for the Taylor coefficients $\hat{q}$, we determine all the multi-indices $k \in K^*$ such that $k^* \in K^*$ and then redefine $\hat{q}_{k^*}$, for each $k^* \neq k$, by setting it equal to $\hat{q}_k$ (if $k^* = k$ we set it equal to $\text{Re}(\hat{q}_k)$). The symmetry implies that $\|x - \hat{x}\|_{X^*} = \|x^* - \hat{x}\|_{X^*}$ for all $x \in X$. Hence $B_r(\hat{x})^* = B_r(\hat{x})$, which motivates the assumption that $x^* \in U$.

Transversality We end this section with a sufficient condition for proving that a connecting orbit is transverse. The key observation is summarized in the following lemma:

Lemma 3.3.21. Suppose $a, \tilde{a} \in \bigoplus_{i=1}^{m} \ell^1_{\nu_i, n}$ are real. Let $u, w : [0, 1] \to \mathbb{R}^n$ denote the maps associated to $a$ and $\tilde{a}$, respectively, i.e.,

$$
u_i := a_0^i + 2 \sum_{k=1}^{\infty} a_k^i T_k^i,$$

$$w_i := \tilde{a}_0^i + 2 \sum_{k=1}^{\infty} \tilde{a}_k^i T_k^i.$$

Then $\tilde{a} \in \ker(DF_u(a))$ if and only if $w'(t) = LDg(u(t)) w(t)$ on $[0, 1]$.

Proof. A straightforward computations shows that $\tilde{a} \in \ker(DF_u(a))$ if and only if

$$ \begin{align*}
    k \tilde{a}_k^i - \frac{L(t_i - t_{i-1})}{4} (Dc_{k-1}(a^i) \tilde{a}_i^i - Dc_{k+1}(a^i) \tilde{a}_i^i) &= 0, \quad 1 \leq i \leq m, \quad k \in \mathbb{N}, \\
    \tilde{a}_0^i - \tilde{a}_0^{i-1} + 2 \sum_{l=1}^{\infty} (-1)^l \tilde{a}_l^i - \tilde{a}_l^{i-1} &= 0, \quad 2 \leq i \leq m.
\end{align*}$$
Furthermore, substitution of the expression in \( (3.17) \) for \( Dc(a^i) \bar{a}^i \) shows that the above system of equations is equivalent to

\[
\begin{align*}
\frac{dw_i}{dt}(t) &= LDg(u_i(t))w_i(t), \quad t \in [t_{i-1}, t_i], \quad 1 \leq i \leq m, \\
w_{i-1}(t_{i-1}) &= w_i(t_{i-1}), \quad 2 \leq i \leq m,
\end{align*}
\]

which proves the statement (see Proposition 3.3.13).

We are now ready to formulate a sufficient criterion for establishing the transversality of a connecting orbit.

**Proposition 3.3.22.** Suppose \( x \in \mathcal{X}_\nu \) is symmetric and \( F(x) = 0 \). If \( DF(x) \) is injective, then \( x \) corresponds to a transverse connecting orbit.

**Proof.** It is shown in Proposition 3.3.19 that \( x \) corresponds to a connecting orbit \( u \) from \( p_0 \) to \( q_0 \) with the property that \( u(0) = P(\theta) \in W^u_{\text{loc}}(p_0) \) and \( u(1) = Q(\phi) \in W^s_{\text{loc}}(q_0) \). To show that \( u \) is transverse, first observe that the mappings

\[
P \circ \iota_u : \mathbb{R}^{\text{sym}, \text{re}}_\nu \subset \mathbb{R}^{n_u} \to \mathbb{R}^n, \quad Q \circ \iota_s : \mathbb{R}^{\text{sym}, \text{re}}_\nu \subset \mathbb{R}^{n_s} \to \mathbb{R}^n,
\]

are parameterizations of \( W^u_{\text{loc}}(p_0) \) and \( W^s_{\text{loc}}(q_0) \), respectively, by Proposition 3.3.3 and Remark 3.3.6. Hence \( \tilde{\theta} \mapsto \varphi(t, P \circ \iota_u(\tilde{\theta})) \), where \( \varphi \) denotes the flow generated by \( Lg \), is a diffeomorphism from \( \mathbb{R}^{\text{sym}, \text{re}}_\nu \) into \( W^u(p_0) \) for any \( t \in \mathbb{R} \). Therefore, its derivative \( D_x\varphi(t, P(\theta))D_\theta P(\theta) \iota_u \) (evaluated at \( \tilde{\theta} = \iota_u^{-1}(\theta) \)) is an isomorphism from \( \mathbb{R}^{n_u} \) onto \( T_{\varphi(t,P(\theta))}W^u(p_0) \). Similarly, \( D_{\theta}Q(\phi) \iota_s \) is an isomorphism from \( \mathbb{R}^{n_s} \) onto \( T_{Q(\phi)}W^s(q_0) \). Consequently, since \( \varphi(1, P(\theta)) = Q(\phi) = u(1) \), the linear map

\[
\Phi_1 := [D_x\varphi(1, P(\theta))D_\theta P(\theta) \iota_u - D_{\phi}Q(\phi) \iota_s]
\]

is a surjection from \( \mathbb{R}^{n_u} \times \mathbb{R}^{n_s} = \mathbb{R}^{n+1} \) onto \( (T_u(1)W^u(p_0) + T_u(1)W^s(q_0)) \subset \mathbb{R}^n \).

Now, suppose \( DF(x) \) is injective but \( u \) is not transverse. Then the intersection of \( W^u(p_0) \) and \( W^s(q_0) \) is (in particular) not transverse at \( u(1) \), since \( u \) is transverse if and only if it is transverse at a point. Hence the map \( \Phi_1 : \mathbb{R}^{n+1} \to \mathbb{R}^n \) cannot be surjective. Therefore, \( \dim(\ker(\Phi_1)) \geq 2 \). Consequently, there exist two linearly independent vectors \( \begin{bmatrix} \tilde{\theta}_1 \\ \varphi_1 \end{bmatrix}, \begin{bmatrix} \tilde{\theta}_2 \\ \varphi_2 \end{bmatrix} \in \ker(\Phi_1) \subset \mathbb{R}^{n_u} \times \mathbb{R}^{n_s} \). We will show that this leads to a contradiction by constructing a nontrivial element in the kernel of \( DF(x) \).

Define \( \xi_1, \xi_2 : [0, 1] \to \mathbb{R}^n \) by

\[
\xi_j(t) := D_x\varphi(t, P(\theta))D_\theta P(\theta) \iota_u \begin{bmatrix} \tilde{\theta}_j \\ \varphi_j \end{bmatrix}, \quad j \in \{1, 2\},
\]

where \( u(1) = Q(\phi) \).
then a straightforward computation shows that

\[
\begin{align*}
\frac{d\xi_j}{dt}(t) &= LDg(u(t))\xi_j(t), \quad t \in [0, 1], \\
\xi_j(0) &= D\theta P(\theta) t_u (\tilde{\theta}_j), \\
\xi_j(1) &= D\phi Q(\phi) t_s (\tilde{\phi}_j),
\end{align*}
\]  

(3.25)

where the boundary condition at \( t = 1 \) follows from the fact that \( \tilde{\theta}_1, \tilde{\phi}_1, \tilde{\theta}_2, \tilde{\phi}_2 \in \ker(\Phi_1) \). Further note that \( \xi_1(t) \) and \( \xi_2(t) \) are linearly independent for each \( t \in [0, 1] \), since the vectors \( \tilde{\theta}_1, \tilde{\phi}_1 \) and \( \tilde{\theta}_2, \tilde{\phi}_2 \) are, and the operators \( D\theta P(\theta) t_u \) and \( D\phi Q(\phi) t_s \) are injective. Consequently, since

\[
\eta = \sum_{i=1}^{m} \mathbf{1}_{[u_{i-1}, u_i]} \left( a^i_0 + 2 \sum_{k=1}^{\infty} a^i_k T^i_k \right), \quad a^i \in \ell^1_{\nu_i, n},
\]

where \( a^1, \ldots, a^m \) are real, there exist (unique) real Chebyshev coefficients \( b_1, b_2 \in \bigoplus_{i=1}^{m} \ell^1_{\nu_i, n} \) such that

\[
\xi_j = \sum_{i=1}^{m} \mathbf{1}_{[u_{i-1}, u_i]} \left( (b^j_i)_0 + 2 \sum_{k=1}^{\infty} (b^j_i)_k T^i_k \right), \quad j \in \{1, 2\}.
\]

In particular, note that any linear combination of \( b_1 \) and \( b_2 \) corresponds to a solution of (3.25) and is thus an element in \( \ker(DF_u(a)) \) by Lemma 3.3.21. Now, set

\[
\lambda^u := 0_{n_u}, \quad \lambda^s := 0_{n_s}, \quad \tilde{p} := 0, \quad \tilde{q} := 0
\]

and

\[
\lambda_j := (t_u (\tilde{\theta}_j), t_s (\tilde{\phi}_j), \lambda^u, \lambda^s, b_j, \tilde{p}, \tilde{q}) \in \mathcal{X}_\nu, \quad j \in \{1, 2\}.
\]

If \( D\eta(a)b_j = 0 \) for some \( j \in \{1, 2\} \), where \( \eta \) is the phase condition defined in (3.24), then a straightforward computation shows that \( h_j \in \ker(DF_u(a)) \). Otherwise, without loss of generality, we may assume that \( D\eta(a)b_1 \neq 0 \) and set

\[
h := h_2 - \frac{D\eta(a)b_2}{D\eta(a)b_1} h_1.
\]

Then \( h \in \ker(DF_u(a)) \) as before by a straightforward computation. In either case, \( h \neq 0 \). Therefore, we have reached a contradiction, since \( DF(x) \) is assumed to be injective. Hence \( u \) must be transverse.

\[ \square \]

**Remark 3.3.23.** In practice, the injectivity of \( DF(x) \) follows directly from our computer-assisted proof (a contraction argument) and is thus obtained for "free", see Remark 3.4.10.

97
3.4 Functional analytic setup

In this section we set up a functional analytic framework for establishing the existence of an isolated zero of \( F \). We start by introducing some notation and a finite dimensional reduction of \( F \). We then combine numerical simulation and analysis on paper to set up a Newton-like operator \( T \) whose fixed points correspond to zeros of \( F \). Finally, we derive a finite number of inequalities to establish that \( T \) is a contraction in a neighborhood of an approximate zero.

3.4.1 Projections

In this section we define projections on both the range and domain. These projections will help structure the calculations in the following sections.

**Projections in \( \mathcal{X}_\nu \)** Write \( x = (\theta, \phi, \lambda^u, \lambda^s, a^1, \ldots, a^m, p, q) \in \mathcal{X}_\nu \). Let \( 1 \leq i \leq m, 1 \leq j \leq n \) and \( k \subset \mathbb{N}_0 \). Define projections \( \Pi^{i}_{a} : \mathcal{X}_\nu \rightarrow \ell^{1}_{\nu, i}, \Pi^{ij}_{a} : \mathcal{X}_\nu \rightarrow \ell^{1}_{\nu, i} \)

onto the Chebyshev coefficients by

\[
\Pi^{i}_{a} (x) := a^{i}, \quad \Pi^{ij}_{a} (x) := [a^{j}]_{i}, \quad (\Pi^{ijk}_{a} (x))_{l} := \begin{cases} [a^{i}]_{j} & l \in k, \\ 0 & \text{otherwise}. \end{cases}
\]

In particular, if \( k \) is a singleton, we identify \( \Pi^{ijk}_{a} (x) \simeq [a^{i}]_{j} \). Similarly, let \( \mathcal{I} \subset \mathbb{N}^{n_u}_0 \), \( \mathcal{J} \subset \mathbb{N}^{n_s}_0 \) and define projections \( \Pi_{P} : \mathcal{X}_\nu \rightarrow W^{1}_{\nu, u, n}, \Pi^{i}_{P} : \mathcal{X}_\nu \rightarrow W^{1}_{\nu, i} \) and \( \Pi_{Q} : \mathcal{X}_\nu \rightarrow W^{1}_{\nu, u, n}, \Pi^{i}_{Q} : \mathcal{X}_\nu \rightarrow W^{1}_{\nu, i} \)

onto the Taylor coefficients of the (un)stable manifolds by

\[
\Pi_{P} (x) := p, \quad \Pi^{i}_{P} (x) := [p]_{i}, \quad (\Pi^{i}_{P} (x))_{k} := \begin{cases} [p]_{i} & k \in \mathcal{I}, \\ 0 & k \notin \mathcal{I}, \end{cases}
\]

\[
\Pi_{Q} (x) := q, \quad \Pi^{i}_{Q} (x) := [q]_{i}, \quad (\Pi^{i}_{Q} (x))_{k} := \begin{cases} [q]_{i} & k \in \mathcal{J}, \\ 0 & k \notin \mathcal{J}. \end{cases}
\]

As before, if \( \mathcal{I} \) and \( \mathcal{J} \) are singletons, we identify \( (\Pi^{ij}_{P} (x)) \simeq [p]_{j} \), \( (\Pi^{ij}_{Q} (x)) \simeq [q]_{j} \). Finally, we define projections \( \Pi^{i}_{\theta}, \Pi^{i}_{\phi}, \Pi^{i}_{\lambda^u}, \Pi^{i}_{\lambda^s} : \mathcal{X}_\nu \rightarrow \mathbb{C} \) by

\[
\Pi^{i}_{\theta} (x) := \theta^{i}, \quad \Pi^{i}_{\phi} (x) := \phi^{i}, \quad 1 \leq j \leq n_u,
\]

\[
\Pi^{i}_{\lambda^u} (x) := \lambda^{i}, \quad \Pi^{i}_{\lambda^s} (x) := \lambda^{i}, \quad 1 \leq j \leq n_s.
\]

**Remark 3.4.1.** In order to keep the notation and number of symbols to a minimum, we have used the symbol \( k \) as a “dummy” index which, depending on the context, can be either an element in \( \mathbb{N}_0, \mathbb{N}^{n_u}_0 \) or \( \mathbb{N}^{n_s}_0 \).
We shall denote the collection of projections onto the components of \( \mathcal{X}_\nu \) by \( \mathbb{P} \), i.e.,

\[
\mathbb{P} := \left\{ \Pi^j_\theta : 1 \leq j \leq n_u \right\} \cup \left\{ \Pi^j_\phi : 1 \leq j \leq n_s \right\} \cup \left\{ \Pi^j_\lambda : 1 \leq j \leq n_u \right\}
\]

\[
\cup \left\{ \Pi^j_\lambda^* : 1 \leq j \leq n_s \right\} \cup \left\{ \Pi^{ij}_\alpha : 1 \leq i \leq m, 1 \leq j \leq n \right\}
\]

\[
\cup \left\{ \Pi^j_p : 1 \leq j \leq n \right\} \cup \left\{ \Pi^j_q : 1 \leq j \leq n \right\} .
\]

Observe that

\[
\| x \|_{\mathcal{X}_\nu} = \max_{\Pi \in \mathbb{P}} \| \Pi(x) \|_{\Pi(\mathcal{X}_\nu)} ,
\]

where \( \| \cdot \|_{\Pi(\mathcal{X}_\nu)} \) denotes the norm on \( \Pi(\mathcal{X}_\nu) \in \{ \mathbb{C}, \ell^1_{\nu}, W^1_{\nu}, W^1_{\nu} \} \). In the future we shall be a bit more sloppy in describing subsets of \( \mathbb{P} \) by omitting the ranges for the components of the projections. For instance, whenever we write \( \{ \Pi^j_\theta, \Pi^j_\phi, \Pi^j_\lambda^*, \Pi^j_\lambda \} \), we mean to say that this set contains all components associated to these projections, i.e., it contains \( \Pi^j_\theta \) for \( 1 \leq j \leq n_u \), \( \Pi^j_\phi \) for \( 1 \leq j \leq n_s \), etc. We will adopt the same convention for the projections into the range which are introduced at the end of this section.

**Galerkin projection into \( \mathcal{X}_\nu \)** Let \( N \in \mathbb{N}^m \) be a given truncation parameter and define operators \( \Pi^i_N : \ell^1_{\nu_i} \rightarrow \ell^1_{\nu_i} \) by

\[
(\Pi^i_N (a^i))_k := \begin{cases} a^i_k, & 0 \leq k \leq N_i - 1, \\ 0_n, & k \geq N_i, \end{cases}, \quad 1 \leq i \leq m.
\]

Similarly, let \( K = (K^u, K^s) \in \mathbb{N}^n_u \times \mathbb{N}^n_s \), set

\[
\mathcal{K}^u := \{ k \in \mathbb{N}^n_u : k_i \leq K^u_i, 1 \leq i \leq n_u \},
\]

\[
\mathcal{K}^s := \{ k \in \mathbb{N}^n_s : k_i \leq K^s_i, 1 \leq i \leq n_s \},
\]

and define \( \Pi_{K^u} : W^1_{\nu_u} \rightarrow W^1_{\nu_u}, \Pi_{K^s} : W^1_{\nu_s} \rightarrow W^1_{\nu_s} \) by

\[
(\Pi_{K^u} (p))_k := \begin{cases} p_k, & k \in \mathcal{K}^u, \\ 0_n, & k \notin \mathcal{K}^u, \end{cases}, \quad (\Pi_{K^s} (q))_k := \begin{cases} q_k, & k \in \mathcal{K}^s, \\ 0_n, & k \notin \mathcal{K}^s, \end{cases}
\]

respectively. Finally, define the Galerkin-projection \( \Pi^{NK}_{\text{dom}} : \mathcal{X}_\nu \rightarrow \mathcal{X}_\nu \) into the domain by

\[
\Pi^{NK}_{\text{dom}} := I_{2(n_u+n_s)} \oplus \bigoplus_{i=1}^{m} \Pi^i_N \oplus \Pi_{K^u} \oplus \Pi_{K^s},
\]

where \( I_{2(n_u+n_s)} \) is the identity on \( \mathbb{C}^{2(n_u+n_s)} \), and set \( \mathcal{X}^{NK} := \Pi^{NK}_{\text{dom}} (\mathcal{X}_\nu) \).
Remark 3.4.2. In practice, we choose an ordering on the set of multi-indices and identify $[\Pi_{K_u}(p)]_j$ with a column vector in $\mathbb{C}^{\prod_{i=1}^{n_u}(K_i^u+1)}$ (and we do the same for the Taylor coefficients associated to the stable manifold). The Chebyshev coefficients $\Pi^i_N(a)$ are identified with a column vector in $\mathbb{C}^{nN_i}$ as well. Altogether, this yields an identification of $\Pi^{NK}_{\text{dom}}(x)$ with a vector in $\mathbb{C}^\kappa$, where

$$\kappa = n \left( \sum_{i=1}^{m} N_i + \prod_{i=1}^{n_u} (K_i^u + 1) + \prod_{i=1}^{n_s} (K_i^s + 1) \right) + 2(n_u + n_s)$$

since $n_u + n_s = n + 1$. In particular, $\mathcal{X}^{NK} \simeq \mathbb{C}^\kappa$.

**Projections in $\mathcal{Y}_\nu$** Recall that the range consists of elements of the form $y = (y_{t_0}, y_{t_m}, y_{\hat{p}_1}, y_{\hat{q}_1}, y_{\eta}, y_{a}, y_{p}, y_{q})$, see Remark 3.3.17. We shall abuse notation and denote the projections onto the Chebyshev and Taylor coefficients in the range in the same way as for the domain. Furthermore, we define projections $\Pi^j_{t_0}, \Pi^j_{t_m}, \Pi^j_{\hat{p}_1}, \Pi^j_{\hat{q}_1}, \Pi_{\eta} : \mathcal{Y}_\nu \to \mathbb{C}$ by

$$\Pi^j_{t_0}(y) : = y_{t_0}, \quad \Pi^j_{t_m}(y) : = y_{t_m}, \quad 1 \leq j \leq n,$$

$$\Pi^j_{\hat{p}_1}(y) : = [y_{\hat{p}_1}]_j, \quad 1 \leq j \leq n_u,$$

$$\Pi^j_{\hat{q}_1}(y) : = [y_{\hat{q}_1}]_j, \quad 1 \leq j \leq n_s,$$

$$\Pi_{\eta}(y) : = y_{\eta}.$$}

The truncation operators on the range are defined in the same way as for the domain. For this reason we shall use the same notation to denote them. There is one slight modification in the projection onto the Chebyshev coefficients associated to the first domain, however, namely we set

$$\left(\Pi^1_N(a^1)\right)_k := \begin{cases} a^1_k, & 1 \leq k \leq N_1 - 1, \\ 0, & k \geq N_1. \end{cases}$$

Finally, we define the Galerkin-projection $\Pi^{NK}_{\text{ran}}$ into the range by

$$\Pi^{NK}_{\text{ran}} := I_{\mathbb{C}^{n+2}} \oplus \Pi^1_N \oplus \bigoplus_{i=2}^{m} \Pi^i_N \oplus \Pi_{K_u} \oplus \Pi_{K_s}$$

and set $\mathcal{Y}^{NK} := \Pi^{NK}_{\text{ran}}(\mathcal{Y}_\nu)$.

**Remark 3.4.3.** Observe that $\mathcal{Y}^{NK} \simeq \mathbb{C}^\kappa$, since $n_u + n_s = n + 1$. Hence $\mathcal{Y}^{NK}$ and $\mathcal{X}^{NK}$ have the same dimension.
3.4.2 An equivalent fixed-point problem

In this section we construct a Newton-like operator and set up an equivalent fixed
point problem. We start by introducing a finite dimensional reduction of the zero
finding problem amenable to numerical computations:

**Definition 3.4.4** (Finite dimensional reduction). The finite dimensional reduc-
tion $F_{NK} : \mathcal{X}^{NK} \to \mathcal{Y}^{NK}$ of the connecting orbit map $F$ is defined by

$$F_{NK} := \Pi_{\text{ran}}^{NK} \circ F|_{\mathcal{X}^{NK}}.$$  

Next, we construct an approximation of $DF(\hat{x})$ and its inverse by combining
numerical computations and analysis on paper. To this end, assume that we have
computed

(A1) an approximate zero $\hat{x} = \left(\hat{\theta}, \hat{\phi}, \hat{\lambda}^u, \hat{\alpha}, \hat{p}, \hat{\varphi}\right) \in \mathcal{X}^{NK}$ of $F_{NK}$ such that

$$\left(\hat{x}\right)^* = \hat{x},$$

(A2) an approximate injective inverse $A_{NK}$ of $DF_{NK}(\hat{x})$.

If the truncation parameters are sufficiently large and the grid is sufficiently fine,
we expect the linear part of the mappings $F_u$, $F_P$ and $F_Q$ to be dominant in a small
neighborhood of the approximate zero. This motivates the following definitions:

**Definition 3.4.5** (Approximate derivative). The approximate derivative $\hat{DF} : \mathcal{X}_\nu \to \mathcal{Y}_\nu$ at $\hat{x}$ is defined by

$$\Pi \hat{DF} h := \Pi(y), \quad \text{for } \Pi \in \{\Pi_{t_0}^i, \Pi_{\tilde{t}_m}^j, \Pi_{\tilde{p}_s}^j, \Pi_{\tilde{q}_s}^j, \Pi_{\eta}\},$$

$$\left[\Pi_a \hat{DF} h\right]_k = \begin{cases}
(\Pi_a(y))_k, & \delta_{i1} \leq k \leq N_i - 1, \\
(k (\Pi_a(h))_k), & k \geq N_i,
\end{cases}, \quad 1 \leq i \leq m,$$

$$\left[\Pi_P \hat{DF} h\right]_k = \begin{cases}
(\Pi_P(y))_k, & k \in \mathcal{K}^u, \\
\langle \hat{\lambda}^u, k \rangle (\Pi_P(h))_k, & k \notin \mathcal{K}^u,
\end{cases},$$

$$\left[\Pi_Q \hat{DF} h\right]_k = \begin{cases}
(\Pi_Q(y))_k, & k \in \mathcal{K}^s, \\
\langle \hat{\lambda}^s, k \rangle (\Pi_Q(h))_k, & k \notin \mathcal{K}^s,
\end{cases},$$

where $y = DF_{NK}(\hat{x}) \Pi_{\text{dom}}^{NK}(h)$. 

101
Definition 3.4.6 (Approximate inverse). The approximate inverse $A : \mathcal{Y}_\nu \to \mathcal{X}_\nu$ of $DF(\hat{x})$ is defined by

$$\Pi Ah := \Pi(x), \quad \text{for } \Pi \in \left\{ \Pi_a^i, \Pi_a^j, \Pi_a^k, \Pi_a^l, \Pi_a^m \right\},$$

$$(\Pi_a^i Ah)_k := \begin{cases} (\Pi_a^i(x))_k, & 0 \leq k \leq N_i - 1, \\ k^{-1}(\Pi_a^i(h))_k, & k \geq N_i, \end{cases}$$

$$(\Pi_P Ah)_k := \begin{cases} (\Pi_P(x))_k, & k \in K^u, \\ \langle \hat{\lambda}^u, k \rangle^{-1}(\Pi_P(h))_k, & k \not\in K^u, \end{cases}$$

$$(\Pi_Q Ah)_k := \begin{cases} (\Pi_Q(x))_k, & k \in K^s, \\ \langle \hat{\lambda}^s, k \rangle^{-1}(\Pi_Q(h))_k, & k \not\in K^s, \end{cases}$$

where $x = A_{NK}\Pi_{\text{ran}}^{NK}(h)$.

Remark 3.4.7. Note that $A$ is injective, since $A_{NK}$ is.

We are now ready to construct a Newton-like operator for $F$ based at the approximate zero:

Definition 3.4.8 (Newton-like operator). The Newton-like operator $T$ for $F$ based at $\hat{x}$ is defined by $T := I - AF$.

A straightforward computation shows that $T$ maps $\mathcal{X}_\nu$ into itself by construction of the approximate inverse $A$. The weights $\hat{\nu}$ are therefore irrelevant. Furthermore, observe that $T(x) = x$ if and only if $F(x) = 0$, since $A$ is injective. We conclude this section with a theorem which can be used to prove that $T$ is a contraction in a neighborhood of $\hat{x}$ by checking a finite number of inequalities. The theorem is based on a parameterized Newton-Kantorovich method and is often referred to as the radii-polynomial approach (see [43] for instance).

Theorem 3.4.9 (Contraction mapping principle with variable radius). Suppose for each $\Pi \in \mathbb{P}$ there exists bounds $Y_{\Pi}, Z_{\Pi}(r) \geq 0$ such that

$$\|\Pi(T(\hat{x}) - \hat{x})\|_{\Pi(\mathcal{X}_\nu)} \leq Y_{\Pi}, \quad (3.26)$$

$$\sup_{v, h \in B_1(0)} \|\Pi DT(\hat{x} + rv) h\|_{\Pi(\mathcal{X}_\nu)} \leq Z_{\Pi}(r), \quad (3.27)$$

where $\|\cdot\|_{\Pi(\mathcal{X}_\nu)}$ denotes the norm on $\Pi(\mathcal{X}_\nu)$. If there exists a radius $\hat{r} > 0$ such that

$$Z_{\Pi}(\hat{r}) \hat{r} + Y_{\Pi} < \hat{r} \quad (3.28)$$

for all $\Pi \in \mathbb{P}$, then $T : B_{\hat{r}}(\hat{x}) \to B_{\hat{r}}(\hat{x})$ is a contraction.
Proof. A proof can be found in [84]. □

Remark 3.4.10. If $T$ is a contraction on $B_r(\hat{x})$, then there exists a unique zero $\tilde{x} \in B_r(\hat{x})$ of $F$. In particular, $(\hat{x})^* \in B_r(\hat{x})$, since $(\hat{x})^* = \hat{x}$ and $\star$ is norm-preserving. Therefore, since $\hat{x}$ is unique in $B_r(\hat{x})$, we have proven the existence of a real connecting orbit by Proposition 3.3.19. Moreover, $DF(\tilde{x})$ is injective, since

$$
\| I - ADF(\tilde{x}) \|_{B(\chi, \chi)} = \| DT(\tilde{x}) \|_{B(\chi, \chi)} < 1
$$

by (3.27) and (3.28). Hence the connecting orbit is transverse by Proposition 3.3.22.

3.5 Bounds for proving contraction

In this section we compute the bounds as stated in Theorem 3.4.9 to prove that $T$ is a contraction in a neighborhood of $\hat{x}$. To compute these bounds, we need to project and perform analysis on the various subspaces of $X_\nu$. Since the analysis for the unstable and stable manifold is the same, we will only write down the arguments in detail for the unstable manifold and simply state the analogous result for the stable manifold. We have aimed to compute the sharpest bounds whenever possible, but there are occasions in which we have chosen to use slightly less optimal bounds when the reduction in computational complexity outweighed the potential loss in accuracy. Moreover, different considerations have also been taken into account in order to keep the exposition of the theory organized.

3.5.1 $Y$-bounds

In this section we compute bounds for the residual

$$
T(\hat{x}) - \hat{x} = -AF(\hat{x})
$$

as stated in Theorem 3.4.9. To this end, observe that

$$
\Pi AF(\hat{x}) = \Pi_{NK} A_{NK} F_{NK}(\hat{x}), \quad \Pi \in \left\{ \Pi^\ell_\nu, \Pi^\ell_\phi, \Pi^\ell_{\lambda^u}, \Pi^\ell_{\lambda^s} \right\}.
$$

Furthermore, for $k \in \mathbb{N}_0$, $1 \leq i \leq m$, $1 \leq j \leq n$, we have that

$$
\Pi^{ijk}_a AF(\hat{x})
$$

$$
= \begin{cases} 
\Pi^{ijk}_a A_{NK} F_{NK}(\hat{x}), & 0 \leq k \leq N_i - 1, \\
-\frac{L(t_i - t_{i-1})}{4k} \left[ c_{k-1}(\hat{a}^i) - c_{k+1}(\hat{a}^i) \right], & N_i \leq k \leq N_\text{g}_j (N_i - 1) + 1, \\
0, & \text{otherwise},
\end{cases}
$$

103
since \([c_k(\hat{a})]_j = 0\) for all \(k \geq N_{g_j}(N_i - 1) + 1\), see the definition of \(c\) in (4.19). Similarly, for \(k \in \mathbb{N}_0^n\),

\[
\Pi_{P}^{jk} \ AF(\hat{x}) = \begin{cases} 
\Pi_{P}^{jk} A_{NK} F_{NK}(\hat{x}), & k \in \mathcal{K}^u, \\
- \langle k, \hat{\lambda}^u \rangle^{-1} [C_k(\hat{p})]_j, & k \in \mathcal{J}_d^j, \\
0, & \text{otherwise},
\end{cases}
\]

where

\[
\mathcal{J}_d^j := \{k \in \mathbb{N}_0^n : k_i \leq N_{g_j}K^u_i, 1 \leq i \leq n_u \} \cap (\mathcal{K}^u)^c,
\]

since \([C_k(\hat{p})]_j = 0\) for \(k \not\in \mathcal{K}^u\), see (3.12).

The above computations show that there are only a finite number of non-vanishing terms in \(AF(\hat{x})\). Therefore, \(AF(\hat{x})\) can be computed with the aid of a computer. It is now a straightforward task to compute the \(Y\)-bounds by taking the appropriate norms of the above expressions:

**Proposition 3.5.1 (Y-bounds).** Let \(1 \leq i \leq m\) and \(1 \leq j \leq n\). The bounds

\[
Y_{\Pi} := ||\Pi A_{NK} F_{NK}(\hat{x})||_{\nu_i}, \quad \Pi \in \{\Pi_i^j, \Pi_d^j, \Pi_{\lambda^u}, \Pi_{\lambda^s}\},
\]

\[
Y_{a}^{ij} := \|\Pi_{a}^{ij} A_{NK} F_{NK}(\hat{x})\|_{\nu_i} + \frac{L(t_i - t_{i-1})}{2} \sum_{k=N_i}^{N_{g_j}(N_i - 1) + 1} \|c_{k-1}(\hat{a}^i) - c_{k+1}(\hat{a}^i)\|_k^{\nu_i},
\]

\[
Y_{P}^{j} := \|\Pi_{P}^{j} A_{NK} F_{NK}(\hat{x})\|_{\nu_u} + \sum_{k \in \mathcal{J}_d^j} ||C_k(\hat{p})||_j \|\langle k, \hat{\lambda}^u \rangle\|^{-1} \nu_u^{|k|},
\]

\[
Y_{Q}^{j} := \|\Pi_{Q}^{j} A_{NK} F_{NK}(\hat{x})\|_{\nu_s} + \sum_{k \in \mathcal{J}_d^j} ||C_k(\hat{q})||_j \|\langle k, \hat{\lambda}^s \rangle\|^{-1} \nu_s^{|k|}
\]

satisfy the estimate in (3.26).

**3.5.2 Z-bounds**

In this section we compute bounds for \(DT\) as stated in Theorem 3.4.9. To this end, let \(r > 0, v, h \in B_1(0)\) be arbitrary and observe that

\[
\begin{align*}
DT(\hat{x} + rv) h &= \left( I - A\hat{D}F \right) h - A \left( DF(\hat{x} + rv) - DF(\hat{x}) + DF(\hat{x}) - \hat{D}F \right) h.
\end{align*}
\]
We shall use this decomposition to compute quadratic polynomials \( Z(\Pi) \) which satisfy the condition in (3.27). Furthermore, throughout this section we shall write
\[
h = \left( \hat{\theta}, \hat{\phi}, \hat{\lambda}^u, \hat{\lambda}^s, \tilde{a}, \tilde{p}, \tilde{q} \right).
\]

Let us start with analyzing the easiest term which measures the quality of the approximate derivative and inverse:

**Lemma 3.5.2.** Let \( \Pi \in \mathbb{P} \), then
\[
\left\| \Pi \left( I - \hat{A} \hat{D} \right) \right\|_{\mathcal{B}(\mathcal{X}_u, \Pi(\mathcal{X}_u))} \leq \left\| \Pi (I_{NK} - A_{NK}DF_{NK}(\hat{x})) \right\|_{\mathcal{B}(\mathcal{X}_u^{NK}, \Pi(\mathcal{X}_u^{NK}))}.
\]

**Proof.** It suffices to observe that
\[
\Pi_{NK} (I - \hat{A} \hat{D}) = I_{NK} - A_{NK}DF_{NK}(\hat{x}),
\]
where \( I_{NK} \) is the identity on \( \mathcal{X}_u^{NK} \). The latter equality holds because \( A \) and \( \hat{D} \) are exact inverses of each other on the subspaces associated to the “tails” in \( W^1_{\nu,u,n} \), \( W^1_{\nu,s,n} \) and \( \bigoplus_{i=1}^m \ell^1_{\nu,s,n} \).

**Remark 3.5.3.** Note that the computation of the stated bound is finite for each \( \Pi \in \mathbb{P} \), since \( I_{NK} - A_{NK}DF_{NK}(\hat{x}) \) is a finite dimensional matrix.

**Chebyshev series: convolution terms**

In this section we develop tools for analyzing the terms

\[
\Pi_{ai}^{ijN} \left( DF(\hat{x}) - \hat{D} \hat{F} \right), \quad 1 \leq i \leq m, \ 1 \leq j \leq n,
\]

which will be used extensively in Section 3.5.2 to compute the \( Z \)-bounds. We start with the observation that

\[
\left( \Pi_{ai}^{i} \left( DF(\hat{x}) - \hat{D} \hat{F} \right) h \right)_k = -\frac{L (t_i - t_{i-1})}{4} \begin{cases} Dc_{k-1} \left( \hat{a}^i \right) \hat{a}_{\infty}^i - Dc_{k+1} \left( \hat{a}^i \right) \hat{a}_{\infty}^i, & 1 \leq k \leq N_i - 1, \\ Dc_{k-1} \left( \hat{a}^i \right) \tilde{a}^i - Dc_{k+1} \left( \hat{a}^i \right) \tilde{a}^i, & k \geq N_i, \end{cases}
\]

for \( 1 \leq i \leq m \), where

\[
(\hat{a}_{\infty}^i)_k := \begin{cases} 0_n, & 0 \leq k \leq N_i - 1, \\ \hat{a}^i_k, & k \geq N_i. \end{cases}
\]

The goal is to construct workable matrix representations for both the finite (truncated) and tail part of (3.29).
To construct suitable matrix representations for (3.29), recall that
\[
D[\epsilon_j (\hat{\alpha}^i) \hat{\alpha}^i = \sum_{l=1}^{n} \hat{g}^{ijl} * [\hat{\alpha}^i]_l, \quad 1 \leq i \leq m, \ 1 \leq j \leq n,
\]
where the coefficients \( \hat{g}^{ijl} \) are defined in (3.18), see Remark 3.3.9. In particular, note that \( \hat{g}^{ijl} = 0 \) for \( k \geq M_{jl} (N_i - 1) + 1 \), where \( M_{jl} \) is the order of \( \partial g_j / \partial x_l \), since \( \hat{a}_k = 0 \) for \( k \geq \tilde{M}_l = M_i (N - 1) + 1 \), where \( M_i, N \in \mathbb{N}, \nu > 1 \), and construct explicit matrix representations for the mappings \( B(a), \Gamma(a) : \ell_1^\nu \to \ell_1^\nu \) defined by
\[
[B(a)\tilde{a}]_k := \begin{cases} (a * \tilde{a})_{k-1} - (a * \tilde{a})_{k+1} & 1 \leq k \leq N - 1, \\ 0 & k = 0 \text{ or } k \geq N, \end{cases}
\]
\[
[\Gamma(a)\tilde{a}]_k := \begin{cases} 0 & 0 \leq k \leq N - 1, \\ (a * \tilde{a})_{k-1} - (a * \tilde{a})_{k+1} & k \geq N, \end{cases}
\]
where
\[
(\tilde{a}_\infty)_k := \begin{cases} 0 & 0 \leq k \leq N - 1, \\ \tilde{a}_k & k \geq N. \end{cases}
\]

Remark 3.5.4. The parameters \( N \) and \( \nu \) in this section are not to be confused with the vector valued ones used throughout this paper. In practice, we set \( a = \hat{g}^{ijl}, N = N_i \) and \( M = M_{jl} \). In particular, observe that
\[
\left[ \Pi^{ij}_{a} \left( DF (\hat{x}) - \overline{DF} \right) h \right]_{k=1}^{N_i-1} = - \frac{L (t_i - t_{i-1})}{4} \left[ \sum_{l=1}^{n} B (\hat{g}^{ijl}) \Pi^{il}_{a} (h) \right]_{k=1}^{N_i-1},
\]
and
\[
\left[ \Pi^{ij}_{a} \left( DF (\hat{x}) - \overline{DF} \right) h \right]_{k=N_i}^{\infty} = - \frac{L (t_i - t_{i-1})}{4} \left[ \sum_{l=1}^{n} \Gamma (\hat{g}^{ijl}) \Pi^{il}_{a} (h) \right]_{k=N_i}^{\infty},
\]
by (3.29).

We begin by extending \( a \) and \( \tilde{a} \) to “bi-infinite” sequences, by setting \( a_{-k} := a_k \) and \( \tilde{a}_{-k} := \tilde{a}_k \) for \( k \in \mathbb{N} \), and constructing a bi-infinite matrix representation for the map
\[
\tilde{a} \mapsto [(a * \tilde{a})_{k-1} - (a * \tilde{a})_{k+1}]_{k \in \mathbb{Z}}.
\]

106
We then convert this bi-infinite matrix representation to an “one-sided” matrix representation by using appropriate reflections, which in turn will be used to construct the desired matrix representations for $B(a)$ and $\Gamma(a)$. To be more precise, first observe that

\[(a * \tilde{a})_{k \in \mathbb{Z}}\]

Here we have identified elements in $\ell^1$ with bi-infinite column vectors (with respect to the ordering as depicted above). The bandwidth of this bi-infinite matrix is $\tilde{M} - 1$, since $a_k$ vanishes for $|k| \geq \tilde{M}$. The shaded regions in grey indicate the position of the “zeroth” row and column. Set $b_k := a_{k-1} - a_{k+1}$ for $-\tilde{M} \leq k \leq \tilde{M}$, then it follows from the above expression that

\[
[(a * \tilde{a})_{k-1} - (a * \tilde{a})_{k+1}]_{k \in \mathbb{Z}}
\]

In particular, this bi-infinite matrix has bandwidth $\tilde{M}$.

Next, we convert the latter matrix representation to an one-sided representation on $\mathbb{N}_0$ by “reflecting” all elements on the left hand-side of the zeroth column to the
right and ignoring the rows with negative indices. This yields

\[
\begin{bmatrix}
(a \ast \tilde{a})_{k-1} - (a \ast \tilde{a})_{k+1}
\end{bmatrix}_{k \in \mathbb{N}_0}
\]

\[
\begin{bmatrix}
\begin{bmatrix}
b_0 & \cdots & b_{-\tilde{M}} \\
\vdots & \ddots & \vdots \\
 b_{\tilde{M}} & \cdots & b_0 & \cdots & b_{-\tilde{M}} \\
 b_{\tilde{M}} & \cdots & b_0 & \cdots & b_{-\tilde{M}} \\
 & \ddots & \vdots & \ddots & \vdots \\
 & & \ddots & \vdots & \end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\tilde{a}_0 \\
\tilde{a}_1 \\
\vdots \\
\tilde{a}_{\tilde{M}}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\begin{bmatrix}
b_1 & \cdots & b_{\tilde{M}} \\
\vdots & \ddots & \vdots \\
 b_{\tilde{M}} & \cdots & b_0 & \cdots & b_{-\tilde{M}} \\
 & \ddots & \vdots & \ddots & \vdots \\
 & & \ddots & \ddots & \vdots \\
 & & & \ddots & \end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\tilde{a}_1 \\
\vdots \\
\tilde{a}_{\tilde{M}}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\begin{bmatrix}
b_0 & \cdots & b_{-\tilde{M}} \\
\vdots & \ddots & \vdots \\
 b_{\tilde{M}} & \cdots & b_0 & \cdots & b_{-\tilde{M}} \\
 b_{\tilde{M}} & \cdots & b_0 & \cdots & b_{-\tilde{M}} \\
 & \ddots & \vdots & \ddots & \vdots \\
 & & \ddots & \vdots & \end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\tilde{a}_0 \\
\tilde{a}_1 \\
\vdots \\
\tilde{a}_{\tilde{M}}
\end{bmatrix}
+ \begin{bmatrix}
\begin{bmatrix}
b_1 & \cdots & b_{\tilde{M}} \\
\vdots & \ddots & \vdots \\
 b_{\tilde{M}} & \cdots & b_0 & \cdots & b_{-\tilde{M}} \\
 & \ddots & \vdots & \ddots & \vdots \\
 & & \ddots & \ddots & \vdots \\
 & & & \ddots & \end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\tilde{a}_1 \\
\vdots \\
\tilde{a}_{\tilde{M}}
\end{bmatrix}
\]

(3.33)

Altogether, the sum of the above two matrices, which we will denote by \(S(a)\), constitutes an infinite dimensional matrix representation of the map

\[
\tilde{a} \mapsto \left[(a \ast \tilde{a})_{k-1} - (a \ast \tilde{a})_{k+1}\right]_{k \in \mathbb{N}_0}.
\]

Let \(\tilde{B}(a) \in \mathbb{C}^{(N-1) \times (\tilde{M}+N)}\) denote the finite dimensional submatrix of \(S(a)\) defined by

\[
\tilde{B}(a) := \left\{ S(a)_{ij} : 1 \leq i \leq N - 1, \ 0 \leq j \leq \tilde{M} + N - 1 \right\}.
\]

Note that we are using the convention that the indexing of the rows and columns start at zero rather than at one. In view of (3.30), set the elements in the columns of \(\tilde{B}(a)\) with index \(0 \leq j \leq N - 1\) to zero and let \(B(a)\) denote the resulting matrix. Finally, let \(\Gamma_\infty(a)\) denote the infinite dimensional matrix which consists of the rows of \(S(a)\) with index \(N\) and higher. Then

\[
[B(a)\tilde{a}]_{k=1}^{N-1} = B(a) \left[\tilde{a}_k\right]_{k=0}^{\tilde{M}+N-1}, \quad [\Gamma(a)\tilde{a}]_{k \geq N} = \Gamma_\infty(a)\tilde{a}
\]

(3.34)

by construction.

In preparation for the analysis in Section 3.5.2, we show that the operator norm of \(\Gamma(a)\) can be computed by considering a sufficiently large finite dimensional submatrix of \(\Gamma_\infty(a)\).
Lemma 3.5.5 (Operator norm of $\Gamma(a)$). Let $\Gamma_N(a) \in \mathbb{C}^{(3\tilde{M} - N+1) \times (2\tilde{M}+1)}$ denote the submatrix $\{ S(a)_{ij} : N \leq i \leq 3\tilde{M}, \ 0 \leq j \leq 2\tilde{M} \}$ of $S(a)$. Then

$$\| \Gamma(a) \|_{\mathcal{B}(\ell_1^\nu, \ell_1^\nu)} = \left\| \begin{bmatrix} 0_{N \times (2\tilde{M}+1)} \\ \Gamma_N(a) \end{bmatrix} \right\|_{\mathcal{B}(\ell_1^\nu, \ell_1^\nu)}.$$  

Proof. It follows directly from the expression in (3.33) that

$$\| \Gamma(a) \xi_k' \|_\nu = \sum_{k=-\tilde{M}}^\tilde{M} |b_k| \nu^k, \ \forall k' \geq 2\tilde{M},$$

where $(\xi_{k'})_{k' \in \mathbb{N}_0}$ are the corner points introduced in Definition 3.2.7. Hence

$$\| \Gamma(a) \|_{\mathcal{B}(\ell_1^\nu, \ell_1^\nu)} = \max_{0 \leq k' \leq 2\tilde{M}} \| \Gamma(a) \xi_k' \|_\nu = \left\| \begin{bmatrix} 0_{N \times (2\tilde{M}+1)} \\ \Gamma_N(a) \end{bmatrix} \right\|_{\mathcal{B}(\ell_1^\nu, \ell_1^\nu)}$$

by Proposition 3.2.9 and the definition of $\Gamma_N(a)$. 

Remark 3.5.6. The latter results shows that the operator norm of $\Gamma(a)$ is determined by its first $2\tilde{M} + 1$ columns.

First order bounds

In this section we compute bounds for

$$A \left( \overline{DF} \left( \hat{x} \right) - \overline{DF} \right) h$$

by projecting it onto all the relevant subspaces of $\mathcal{X}_\nu$. For notational convenience, we shall write $y = \left( \overline{DF} \left( \hat{x} \right) - \overline{DF} \right) h$ throughout this section. We start by computing the difference between the exact and approximate derivative.

A straightforward computation shows that

$$\Pi_{\ell_0}(y) = 2 \sum_{k=N_1}^\infty (-1)^k \tilde{a}_k^l - \sum_{k \notin K^u} \tilde{p}_k \hat{\phi}^k, \quad \Pi_{\ell_m}(y) = 2 \sum_{k=N_m}^\infty \tilde{a}_k^m - \sum_{k \notin K^s} \tilde{q}_k \hat{\phi}^k.$$  

Furthermore, $\Pi(y) = 0$ for $\Pi \in \{ \Pi_{\bar{p}_1}, \Pi_{\bar{q}_1}, \Pi_{\eta} \}$, since the equations associated to $\Pi F$ are linear and only depend on elements in the finite dimensional subspace $\mathcal{X}^{NK}$. Next, set

$$k_{i,j,l} := \left[ 0, (M_{jl} + 1) (N_i - 1) + 1 \right] \cap \mathbb{N}_0, \quad 1 \leq i \leq m, \quad 1 \leq j, l \leq n,$$
where \( M_{jl} \) = order \( \frac{\partial g_j}{\partial x_l} \), then it follows from (3.31), (3.32) and (3.34) that

\[
(\Pi_i^j(y))_k = \begin{cases} 
\sum_{l=N_i}^{\infty} (-1)^l [\tilde{a}_l]_j - \sum_{l=N_i-1}^{\infty} [\tilde{a}_l^{-1}]_j, & k = 0, \\
-L \left( t_i - t_i - 1 \right) \frac{n}{4} \sum_{l=1}^{n} B \left( \hat{g}^{ijl} \right) \Pi_{\tilde{a}^n}^{iilkijl} (h) \right), & 1 \leq k \leq N_i - 1, \\
-L \left( t_i - t_i - 1 \right) \frac{n}{4} \sum_{l=1}^{n} \Gamma \left( \hat{g}^{ijl} \right) \Pi_{\tilde{a}^n}^{il} (h) \right), & k \geq N_i,
\end{cases}
\]

for \( 2 \leq i \leq m \), see Section 3.5.2. Here the matrix-vector product \( B \left( \hat{g}^{ijl} \right) \Pi_{\tilde{a}^n}^{iilkijl} (h) \) is interpreted by using the identification

\[
\Pi_{\tilde{a}^n}^{iilkijl} (h) \simeq \left[ \left[ \tilde{a}_0^i \right]_l \cdots \left[ \tilde{a}_{\max(k_{ijl})}^i \right]_l \right]^T.
\]

The same formula holds for \( i = 1 \) and \( k \in \mathbb{N} \). In particular, if \( i = 1 \), then there is no component to consider for \( k = 0 \). Finally, we compute that

\[
(\Pi_P(y))_k = \begin{cases} 
0_n, & k \in K^u, \\
-DC_k (\hat{p}) \hat{p}, & k \notin K^u,
\end{cases}
\]

(3.38)

Altogether, the above formulae give rise to the decomposition

\[
y = \sum_{j=1}^{n} \left( \Pi_{\tilde{a}^n}^{i} (y) + \Pi_{\tilde{a}^n}^{i} (y) + \sum_{i=2}^{m} \Pi_{\tilde{a}^n}^{ij0} (y) + \sum_{i=1}^{m} \left[ \Pi_{\tilde{a}^n}^{ij1: N_i - 1} (y) + \Pi_{\tilde{a}^n}^{ij N_i: \infty} (y) \right] \\
+ \Pi_{\tilde{a}^n}^{i c} (y) + \Pi_{\tilde{a}^n}^{i c} (y) \right),
\]

(3.39)

where we have set

\[
K^u_c := N^u \setminus K^u, \quad K^s_c := N^s \setminus K^s,
\]

\[
[1 : N_i - 1] : = [1, N_i - 1] \cap \mathbb{N}, \quad [N_i : \infty) := [N_i, \infty) \cap \mathbb{N}.
\]

The strategy is to compute bounds for (3.35) by individually composing each term in the above decomposition with \( A \) and analyzing the associated projections into the domain.

**Remark 3.5.7.** Observe that \( A \Pi_{\tilde{a}^n}^{i j N_i: \infty} \), \( A \Pi_{\tilde{a}^n}^{i c} \) and \( A \Pi_{\tilde{a}^n}^{i c} \), are “diagonal” and “uncoupled” in the sense that the only nonzero projections into the domain are \( \Pi_{\tilde{a}^n}^{i j A \Pi_{\tilde{a}^n}^{i j N_i: \infty}} \), \( \Pi_{\tilde{a}^n}^{j A \Pi_{\tilde{a}^n}^{i j N_i: \infty}} \) and \( \Pi_{\tilde{a}^n}^{j A \Pi_{\tilde{a}^n}^{i j N_i: \infty}} \).
Boundary conditions  We start by considering the terms associated to $A\Pi_{t_0}^j(y)$, $A\Pi_{t_m}^j(y)$ and $A\Pi_{t_0}^{j_0}(y)$, which are related to the boundary conditions.

Lemma 3.5.8. Let $\Pi \in \mathbb{P}$, $2 \leq i \leq m$ and $1 \leq j \leq n$, then

$$\|\Pi A\Pi_{t_0}^j(y)\|_{\Pi(\mathcal{X}_\nu)} \leq \|\Pi A_N K \Pi_{t_0}^j\|_{\Pi(\mathcal{X}_\nu)} \left( \nu_1^{-N_1} + \max_{1 \leq l \leq n_u} \left| \frac{\hat{\theta}_l}{\nu_u} \right|^{K_{l}^{i+1}} \right),$$

(3.40)

$$\|\Pi A\Pi_{t_m}^j(y)\|_{\Pi(\mathcal{X}_\nu)} \leq \|\Pi A_N K \Pi_{t_m}^j\|_{\Pi(\mathcal{X}_\nu)} \left( \nu_m^{-N_m} + \max_{1 \leq l \leq n_s} \left| \frac{\hat{\phi}_l}{\nu_s} \right|^{K_{l}^{i+1}} \right),$$

(3.41)

$$\|\Pi A\Pi_{t_0}^{j_0}(y)\|_{\Pi(\mathcal{X}_\nu)} \leq \|\Pi A_N K \Pi_{t_0}^{j_0}\|_{\Pi(\mathcal{X}_\nu)} \left( \nu_{t_0}^{-N_{t_0}} + \nu_{t_0}^{-N_{t_0-1}} \right),$$

(3.42)

where $\|\cdot\|_{\Pi(\mathcal{X}_\nu)}$ denotes the corresponding norm on $\Pi(\mathcal{X}_\nu)$.

Proof. First observe that

$$\|a^i \mapsto 2 \sum_{k=N_i}^{\infty} a_k^i\|_{\mathcal{B}(\mathcal{E}_{\nu_i}, \mathcal{C})} = \|a^i \mapsto 2 \sum_{k=N_i}^{\infty} (-1)^k a_k^i\|_{\mathcal{B}(\mathcal{E}_{\nu_i}, \mathcal{C})} = \nu_1^{-N_i}, \quad 1 \leq i \leq m,$$

by Proposition 4.2.5. Similarly,

$$\|p \mapsto \sum_{k \in K_{\nu}} p_k \theta^k\|_{\mathcal{B}(W_{\nu}^1, \mathcal{C})} = \sup \left\{ \left| \frac{\hat{\theta}}{\nu_u} \right|^k : k \in \mathbb{N}_{0}^{n_u}, \exists 1 \leq l \leq n_u \text{ such that } k_l \geq K_{l}^{i+1} \right\}$$

$$\leq \max_{1 \leq l \leq n_u} \left| \frac{\hat{\theta}_l}{\nu_u} \right|^{K_{l}^{i+1}}$$

by Proposition 3.2.12, where in the last line we used that $\hat{\theta} \in \mathbb{B}_{\nu_u}$. Finally, the above bounds and the expressions in (3.52) and (3.37) show that

$$\|\Pi_{t_0}^j(y)\| \leq \nu_1^{-N_1} + \max_{1 \leq l \leq n_u} \left| \frac{\hat{\theta}_l}{\nu_u} \right|^{K_{l}^{i+1}}, \quad \|\Pi_{t_m}^j(y)\| \leq \nu_m^{-N_m} + \max_{1 \leq l \leq n_s} \left| \frac{\hat{\phi}_l}{\nu_s} \right|^{K_{l}^{i+1}},$$

$$\|\Pi_{t_0}^{j_0}(y)\| \leq \nu_{t_0}^{-N_{t_0}} + \nu_{t_0}^{-N_{t_0-1}},$$

which proves the statement.

Remark 3.5.9. The computation of the stated bounds is finite for each $\Pi \in \mathbb{P}$, since $A_{NK}$ is a finite dimensional matrix.
Chebyshev coefficients Next, we consider the terms associated to $A\Pi_{a}^{\tilde{i}jN}(y)$. We start with the observation that

$$A\Pi_{a}^{i:j[1:N_{i}-1]}(y) = -\frac{L}{4}(t_{i} - t_{i-1}) \sum_{l=1}^{n} A_{NK}^{i:j[1:N_{i}-1]}B(\tilde{g}_{ijl}) \Pi_{a}^{N_{i}k_{ijl}}(h) \quad (3.43)$$

by (3.37). Note that $A_{NK}^{i:j[1:N_{i}-1]}B(\tilde{g}_{ijl})$ is a finite dimensional matrix which can be explicitly computed on a computer. In particular,

$$\Pi A_{NK}^{i:j[1:N_{i}-1]}B(\tilde{g}_{ijl}), \quad \Pi \in \mathbb{P},$$

corresponds to a finite dimensional matrix representation of a linear operator on $\ell_{\nu}^{1}$. Hence the computation of its operator norm is finite.

**Lemma 3.5.10** (Scalar and Taylor projections). Let $1 \leq i \leq m$, $1 \leq j \leq n$ and $\Pi \in \{\Pi_{a}^{d}, \Pi_{a}^{\phi}, \Pi_{a}^{\lambda u}, \Pi_{a}^{\lambda s}, \Pi_{a}^{\lambda}, \Pi_{a}^{Q}\}$, then

$$\|\Pi A \Pi_{a}^{\tilde{i}jN}(y)\|_{\Pi(x_{\nu})} \leq \frac{L}{4}(t_{i} - t_{i-1}) \sum_{l=1}^{n} \|\Pi A_{NK}^{i:j[1:N_{i}-1]}B(\tilde{g}_{ijl})\|_{B(\ell_{\nu}^{1}, \Pi(x_{\nu}))}.$$  

**Proof.** It suffices to observe that

$$\Pi A \Pi_{a}^{\tilde{i}jN}(y) = \Pi A_{NK}^{i:j[1:N_{i}-1]}, \quad \Pi \in \{\Pi_{a}^{j}, \Pi_{a}^{j}, \Pi_{a}^{j}, \Pi_{a}^{j}, \Pi_{a}^{j}, \Pi_{a}^{j}\},$$

by construction of the approximate inverse $A$. Hence the result follows directly from (3.43). \qed

To analyze the terms $\Pi_{a}^{i:j}A\Pi_{a}^{\tilde{i}jN}(y)$ for $1 \leq i \leq m$, $1 \leq j \leq n$, we first derive a more explicit expression for the tail $A\Pi_{a}^{i:j[1:N_{i}:\infty]}(y)$. For this purpose, define a (infinite dimensional) diagonal matrix $D_{i\infty}^{i}$ by

$$D_{i\infty}^{i} := \begin{bmatrix}
1 & & & \\
\frac{1}{N_{i}} & 1 & & \\
& \frac{1}{N_{i} + 1} & & \\
& & \ddots & \\
& & & 1
\end{bmatrix}, \quad 1 \leq i \leq m.$$

Then it follows from (3.37) and the definition of the approximate inverse that

$$\Pi_{a}^{i:j[1:N_{i}:\infty]}A\Pi_{a}^{\tilde{i}j[1:N_{i}:\infty]}(y)$$

$$= \begin{cases}
0, & (i, j) \neq (i, j), \\
-\frac{L}{4}(t_{i} - t_{i-1}) \sum_{l=1}^{n} D_{i\infty}^{i} \Gamma_{\infty}(\tilde{g}_{ijl}) \Pi_{a}^{i:l}(h), & (i, j) = (i, j).
\end{cases}$$

112
Altogether, by combining the latter result with \((3.43)\), we conclude that if \((i, j) \neq (i, j)\), then
\[
\Pi_{i}^{ij} A \Pi_{i}^{jN}(y) = -
L \frac{(t_{i} - t_{i-1})}{4} \sum_{l=1}^{n} \Pi_{i}^{ij[1:N_{i}-1]} A_{NK} \Pi_{i}^{j[l:1:N_{i}-1]} B (\hat{g}^{ijl}) \Pi_{i}^{l[k;ijl]} (h).
\]

\(3.44\)

Otherwise, if \((i, j) = (i, j)\), then
\[
\Pi_{i}^{ij} A \Pi_{i}^{jN}(y) = -
L \frac{(t_{i} - t_{i-1})}{4} \sum_{l=1}^{n} \Pi_{i}^{ij[1:N_{i}-1]} A_{NK} \Pi_{i}^{j[l:1:N_{i}-1]} B (\hat{g}^{ijl}) D_{i}^{i} (\hat{g}^{ijl}) \Pi_{i}^{l[k;ijl]} (h).
\]

\(3.45\)

We are now ready to compute the desired bounds. Before we proceed, observe that the operator norms of the infinite dimensional matrices in \((3.45)\) can be computed by considering sufficiently large finite dimensional submatrices by the same reasoning as in Lemma 3.5.5. The details are given in the lemma below.

**Lemma 3.5.11** (Chebyshev projections). Let \(1 \leq i, j \leq m\), \(1 \leq i, j \leq n\) and set \(L_{i} := \frac{L(t_{i} - t_{i-1})}{4}\). If \((i, j) \neq (i, j)\), then
\[
\| \Pi_{i}^{ij} A \Pi_{i}^{jN}(y) \|_{\nu_{i}} \leq L_{i} \sum_{l=1}^{n} \| \Pi_{i}^{ij[1:N_{i}-1]} A_{NK} \Pi_{i}^{j[l:1:N_{i}-1]} B (\hat{g}^{ijl}) \|_{B(\ell_{1}^{1}, \ell_{1}^{1})}.
\]

Otherwise, if \((i, j) = (i, j)\), then
\[
\| \Pi_{i}^{ij} A \Pi_{i}^{jN}(y) \|_{\nu_{i}} \leq L_{i} \sum_{l=1}^{n} \| \Pi_{i}^{ij[1:N_{i}-1]} A_{NK} \Pi_{i}^{j[l:1:N_{i}-1]} B (\hat{g}^{ijl}) D_{i}^{i} (\hat{g}^{ijl}) \|_{B(\ell_{1}^{1}, \ell_{1}^{1})},
\]

where
\[
D_{i}^{ijl} := \begin{bmatrix}
\frac{1}{N_{i}} \\
\vdots \\
\frac{1}{3M_{ijl}}
\end{bmatrix}, \quad M_{ijl} := M_{jl} (N_{i} - 1) + 1,
\]

and \(\Gamma_{N} (\hat{g}^{ijl})\) is defined in Lemma 3.5.5.
Lemma 3.5.12. \textit{Let} and \textit{Π} be particular, recall that the only nonzero projections in this case are \( p \) and \( \Pi \). Suppose \( \hat{\ell} \) is the operator norm of \((3.46)\) is completely determined by its columns with index \( 0 \leq k' \leq 2 \tilde{M}_{ijl} \). The corresponding submatrix is given by

\[
\left[ \Pi_{\bar{a}}^{ij[0:N-1]} A_{N,K} \Pi_{\bar{a}}^{ij[1:N-1]} B \left( \hat{g}^{ijl} \right) 0_{N \times \infty} \right]_{[\tilde{a}^*]_l}
\]

A similar computation as in Lemma 3.5.5 shows that

\[
\left\| \left[ \Pi_{\bar{a}}^{ij[0:N-1]} A_{N,K} \Pi_{\bar{a}}^{ij[1:N-1]} B \left( \hat{g}^{ijl} \right) 0_{N \times \infty} \right] \xi \right\|_{\nu_i}
\]

\[
= \sum_{k=0}^{\tilde{M}_{ijl}} \left| \hat{g}_{k-1}^{ijl} - \hat{g}_{k+1}^{ijl} \right| \frac{\nu_i^k}{k + 2 \tilde{M}_{ijl} + k'}, \quad \forall k' \geq 2 \tilde{M}_{ijl},
\]

where \( (\xi_{k'})_{k' \in \mathbb{N}_0} \) are the corner points introduced in Definition 3.2.7. Note that the latter quantity is decreasing for \( k' \geq 2 \tilde{M}_{ijl} \). Hence, by Proposition 3.2.9, the operator norm of \((3.46)\) is completely determined by its columns with index \( 0 \leq k' \leq 2 \tilde{M}_{ijl} \). Therefore, the result now follows from \((3.45)\) and the triangle inequality. \( \Box \)

\textbf{Taylor coefficients} Finally, we consider the terms \( A \Pi_{P}^{\mathcal{K}^u} (y) \) and \( A \Pi_{Q}^{\mathcal{K}^u} (y) \). In particular, recall that the only nonzero projections in this case are \( \Pi_{P}^{\mathcal{K}^u} (y) \) and \( A \Pi_{Q}^{\mathcal{K}^u} (y) \). To study these terms, we will use the following result:

Lemma 3.5.12. \textit{Let} \( M \in \mathbb{N} \) and set \( \mathcal{K}^u_M := \{ k \in \mathbb{N}_0^u : k_i \leq MK^u_i, 1 \leq i \leq n_u \} \). Suppose \( p \in W_{\nu_u}^1 \) satisfies \( p_k = 0 \) whenever \( k \notin \mathcal{K}^u_M \) and define an operator \( \Lambda_u(p) : W_{\nu_u}^1 \rightarrow W_{\nu_u}^1 \) by

\[
(\Lambda_u(p)w)_k := \begin{cases} 0, & k \in \mathcal{K}^u, \\
 \left( k, \lambda^u \right)^{-1} (p \ast w)_k, & k \notin \mathcal{K}^u. \end{cases}
\]

Then \( \Lambda_u(p) \) is bounded and

\[
\| \Lambda_u(p) \|_{B(W_{\nu_u}^1, W_{\nu_u}^1)} \leq \max_{1 \leq j \leq n_u} \left\{ \sum_{k \in \mathcal{K}^u_M, k_j \geq K_j + 1 - l_j} \left| p_k \right| \nu_u^{\left| k \right|} \left( k_j + l_j \right) \left| \text{Re} \left( \hat{\lambda}^u_j \right) \right| + \sum_{1 \leq i \leq n_u} \left| k_i \right| \left| \text{Re} \left( \hat{\lambda}^u_i \right) \right| \right\}_{l_j=0}^{K_j + 1}.
\]
Proof. It follows directly from the Banach algebra estimate that $\Lambda^u(p)$ is bounded. To obtain the stated bound for operator norm, first note that
\[
(p \ast w)_k = \sum_{\alpha + \beta = k} p_{\alpha \beta} w_{\beta} = \sum_{\max\{0, k_i - MK^u_i\} \leq \beta_i \leq k_i} p_{k-\beta} w_{\beta},
\]
for any $k \in \mathbb{N}_0^n$. In particular,
\[
(p \ast \xi)_k = \begin{cases} p_{k-\nu u}^{-|l|}, & \max\{0, k_i - MK^u_i\} \leq l_i \leq k_i, \ 1 \leq i \leq n_u, \\ 0, & \text{otherwise}, \end{cases}
\]
for any $l \in \mathbb{N}_0^n$, where $(\xi)_l \in \mathbb{N}_0^n$ are the corner points introduced in Definition 3.2.11. Consequently,
\[
\|\Lambda^u(p)\xi\|_{\nu u} = \sum_{k \in \mathbb{K}^u_{\nu u} \cap (l + \mathbb{K}^u_M)} |p_{k-l}| \left| \left\langle k, \hat{\lambda}^u \right\rangle \right|^{-1} \nu_u^{k-|l|} = \sum_{k \in \mathbb{K}^u_{\nu u} \cap (l + \mathbb{K}^u_M)} |p_k| \left| \left\langle k + l, \hat{\lambda}^u \right\rangle \right|^{-1} \nu_u^{k}. \tag{3.47}
\]

Next, observe that for any fixed $l \in \mathbb{N}_0^n$ and $1 \leq j \leq n_u$,
\[
\left| \left\langle k + l, \hat{\lambda}^u \right\rangle \right| = \left| \sum_{i=1}^{n_u} (k_i + l_i) \hat{\lambda}^u_i \right| \geq \left| \sum_{i=1}^{n_u} (k_i + l_i) \text{Re} \left( \hat{\lambda}^u_i \right) \right| \geq (k_j + l_j) \left| \text{Re} \left( \hat{\lambda}^u_j \right) \right| + \sum_{1 \leq i \leq n_u, i \neq j} k_i \left| \text{Re} \left( \hat{\lambda}^u_i \right) \right|,
\]
where in the last line we used the fact that all the $\hat{\lambda}^u_i$ have the same sign. Therefore, the term in (3.47) is bounded by
\[
\max_{1 \leq j \leq n_u} \sum_{k \in \mathbb{K}^u_M} |p_{k-j} \nu_u^{k-j}| \left| \left\langle k_j + l_j, \hat{\lambda}^u_j \right\rangle \right| + \sum_{1 \leq i \leq n_u, i \neq j} k_i \left| \text{Re} \left( \hat{\lambda}^u_i \right) \right|.
\]
Finally, note that for any fixed $1 \leq j \leq n_u$, the above sum is strictly decreasing for $l_j \geq K^u_j + 1$. Hence the desired result now follows from Proposition 3.2.12. \qed
Remark 3.5.13. A similar statement holds for the map associated to the stable manifold. The corresponding operator (for \( q \in W^1_{\nu_s} \)) is denoted by \( \Lambda_s(q) : W^1_{\nu_s} \to W^1_{\nu_s} \).

We are now ready to compute the required bounds. To this end, observe that

\[
D [C]_j (\tilde{p}) \tilde{p} = \sum_{l=1}^{n} \hat{G}^{j[l]}_u \ast [\tilde{p}]_l, \quad 1 \leq j \leq n,
\]

by the reasoning in Remark 3.3.9, where \([C]_j\) denotes the \( j \)-th component of \( C \) and \( \hat{G}^{j[l]}_u \) are the Taylor coefficients of \( \frac{\partial g_j}{\partial x_l} (x \mapsto \sum_{k \in K} \hat{p}_k x^k) \), \( 1 \leq l \leq n \).

The coefficients \( \hat{G}^{j[l]}_u \) associated to the stable manifold are defined similarly.

**Lemma 3.5.14 (Projection onto the Taylor coefficients).** Let \( 1 \leq j \leq n \), then

\[
\left\| \Pi_P^{j} A \Pi^{j_K^c} (y) \right\|_{W^1_{\nu_u}} \leq \sum_{l=1}^{n} \left\| \Lambda_u \left( \hat{G}^{j[l]}_u \right) \right\|_{B(W^1_{\nu_u},W^1_{\nu_u})},
\]

(3.49)

\[
\left\| \Pi_Q^{j} A \Pi^{j_K^c} (y) \right\|_{W^1_{\nu_s}} \leq \sum_{l=1}^{n} \left\| \Lambda_s \left( \hat{G}^{j[l]}_s \right) \right\|_{B(W^1_{\nu_s},W^1_{\nu_s})}.
\]

(3.50)

**Proof.** It suffices to observe that

\[
\Pi_P^{j} A \Pi^{j_K^c} (y) = \sum_{l=1}^{n} \Lambda_u \left( \hat{G}^{j[l]}_u \right) [\tilde{p}]_l
\]

by (3.38), (3.48) and the definition of the approximate inverse. \( \square \)

**First order coefficients of \( Z_\Pi (r) \)** We are now ready to construct the first order terms of the (quadratic) polynomials \( Z_\Pi (r) \) for \( \Pi \in \mathbb{P} \). For this purpose, we first introduce some additional notation. We will denote the bounds in Lemma 3.5.10 and 3.5.11, which measure the quality of the approximate derivative and inverse, by \( H_\Pi \). The bounds in (3.40), (3.41) and (3.42), which are related to the boundary conditions, will be denoted by \( Z^{1, j_0}_\Pi \), \( Z^{1, j_m}_\Pi \) and \( Z^{1, j_0}_\Pi \), respectively. The bounds in Lemmas 3.5.10 and 3.5.11, which are related to the differential equation, will be denoted by \( Z^{1, j_1}_\Pi [1:N_{i-1}] \) and \( Z^{1, j_1}_\Pi \), respectively. Finally, the bounds in (3.49) and (3.50), which are related to the invariance equation for the charts on the (un)stable manifolds, will be denoted by \( Z^{1, j}_P \) and \( Z^{1, j}_Q \), respectively.

With the above notation in place, we define

\[
Z^{1}_\Pi := H_\Pi + \sum_{j=1}^{n} \left( Z^{1, j_0}_\Pi + Z^{1, j_m}_\Pi + \sum_{i=2}^{m} Z^{1, j_0}_\Pi + \sum_{i=1}^{m} Z^{1, j_1}[1:N_{i-1}] \right),
\]

116
for \( \Pi \in \{ \Pi_{\hat{\phi}}^j, \Pi_{\phi}^j, \Pi_{\lambda^a}^j, \Pi_{\lambda^s}^j \} \) and

\[
Z_{\Pi_{\lambda^a}^j}^1 := H_{\Pi_{\lambda^a}^j} + \sum_{j=1}^{n} \left( Z_{\Pi_{\lambda^a}^j}^{1,j_0} + Z_{\Pi_{\lambda^a}^j}^{1,j_m} + \sum_{i=2}^{m} \left( Z_{\Pi_{\lambda^a}^j}^{1,i_j0} + \sum_{i=1}^{m} Z_{\Pi_{\lambda^a}^j}^{1,i_jN_j} \right) \right),
\]

\[
Z_{\Pi_{\phi}^j}^1 := H_{\Pi_{\phi}^j} + \sum_{j=1}^{n} \left( Z_{\Pi_{\phi}^j}^{1,j_0} + Z_{\Pi_{\phi}^j}^{1,j_m} + \sum_{i=2}^{m} \left( Z_{\Pi_{\phi}^j}^{1,i_j0} + \sum_{i=1}^{m} Z_{\Pi_{\phi}^j}^{1,i_j1:N_{1-1}} \right) + Z_{\Pi_{\phi}^j}^{1,j} \right),
\]

\[
Z_{\Pi_{\lambda^a}^j}^1 := H_{\Pi_{\lambda^a}^j} + \sum_{j=1}^{n} \left( Z_{\Pi_{\lambda^a}^j}^{1,j_0} + Z_{\Pi_{\lambda^a}^j}^{1,j_m} + \sum_{i=2}^{m} \left( Z_{\Pi_{\lambda^a}^j}^{1,i_j0} + \sum_{i=1}^{m} Z_{\Pi_{\lambda^a}^j}^{1,i_j1:N_{1-1}} \right) + Z_{\Pi_{\lambda^a}^j}^{1,j} \right),
\]

for \( 1 \leq i \leq m, \ 1 \leq j \leq n \).

**Second order bounds**

In this section we compute bounds for

\[
\left. A \left( DF (\hat{x} + rv) - DF (\hat{x}) \right) \right|_h. \tag{3.51}
\]

We will compute the desired bounds by projecting (3.51) into the relevant subspaces of \( X_{\nu} \) as in the previous section. We start with the observation that

\[
A \left( DF (\hat{x} + rv) - DF (\hat{x}) \right) h = \int_{0}^{1} AD^2 F (\hat{x} + \tau rv) [v, h] \, d\tau \, r
\]

by the (generalized) Mean-Value Theorem. For notational convenience, we shall write \( y (\tau) = D^2 F (\hat{x} + \tau rv) [v, h] \) and \( v = (\dot{\theta}, \dot{\phi}, \dot{\lambda}, \dot{x}, \dot{a}, \dot{p}, \dot{q}) \). Furthermore, we will denote the max-norm on both \( \mathbb{C}^{n_u} \) and \( \mathbb{C}^{n_v} \) by \( \| \cdot \|_\infty \).

Observe that \( \Pi (y (\tau)) \) hold for \( \Pi \in \{ \Pi_{\hat{\phi}}, \Pi_{\phi}, \Pi_{\lambda^a}, \Pi_{\lambda^s} \} \), since the equations associated to these projections are linear. Furthermore, a straightforward computation shows that

\[
\Pi_{\hat{\phi}}^j (y (\tau)) = - \sum_{i=1}^{n_u} \sum_{k \in \mathbb{N}_0^{n_u}} k_i \left( \dot{\theta} + \tau r \dot{\theta} \right)^{k-e_i} \left[ \hat{p}_k \dot{\theta}_i + \hat{\theta}_k \dot{\theta}_i \right]
\]

\[
- \sum_{i,l=1}^{n_u} \sum_{k \in \mathbb{N}_0^{n_u}} k_i (k_i - \delta_{il}) \left[ \hat{p}_k + \tau r \dot{p}_k \right] \left( \dot{\theta} + \tau r \dot{\theta} \right)^{k-e_i-e_l} \dot{\theta}_i \dot{\theta}_i \tag{3.52}
\]

for \( 1 \leq j \leq n \). An analogous formula holds for \( \Pi_{\phi}^j (y (\tau)) \). Next, let \( 2 \leq i \leq m \),
\[(\Pi_a^i (y (\tau)))_k = \begin{cases} 0, & k = 0, \\ \frac{-L (t_i - t_{i-1})}{4} \left( D^2 c_{k-1} (\dot{a}^i + \tau r \dot{a}^i) \right. & \\ & \left. \frac{-D^2 c_{k+1} (\dot{a}^i + \tau r \dot{a}^i)}{4} \right], \\ k \in \mathbb{N} \end{cases} \]  

(3.53)

This formula is also valid for \( i = 1 \) and \( k \in \mathbb{N} \), but in this case there is no component to consider for \( k = 0 \). Finally, observe that

\[(\Pi_Q (y (\tau)))_k = \begin{cases} D^2 g (\dot{p}_0 + \tau r \dot{p}_0) [\dot{p}_0, \ddot{p}_0], & k = 0, \\ D^3 g (\dot{p}_0 + \tau r \dot{p}_0) [\dot{p}_0, \ddot{p}_0, \dddot{p}_0 + \tau r \dddot{p}_0] + D^2 g (\dot{p}_0 + \tau r \dot{p}_0) [\dddot{p}_0, \dddot{p}_0 + \tau r \dddot{p}_0] - \left( \dddot{\lambda}_u \dddot{p}_0 + \dddot{\lambda}_u \dddot{p}_0 \right), & k = e_i, 1 \leq i \leq n_u, \\ \left( \dddot{\lambda}_u, k \right) \dddot{p}_k + \left( \dddot{\lambda}_u, k \right) \dddot{p}_k - D^2 C_k (\dddot{p} + \tau r \dddot{p}) [\dddot{p}, \dddot{p}], & |k| \geq 2. \end{cases} \]  

(3.54)

The formula for \( \Pi_Q (y (\tau)) \) is analogous.

Altogether, the above formulae give rise to the decomposition

\[ y (\tau) = \sum_{j=1}^{n} \left( \Pi_{t_0}^j (y (\tau)) + \Pi_{t_m}^j (y (\tau)) + \sum_{i=1}^{m} \Pi_{a_i}^{jN} (y (\tau)) + \Pi_{p}^j (y (\tau)) + \Pi_{Q}^j (y (\tau)) \right). \]

We will now follow the same strategy as in the previous section and compute bounds for (3.51) by individually composing each term in the above decomposition with \( A \) and analyzing the associated projections into the domain.

**Boundary conditions** We start by considering the terms associated to

\[ \int_0^1 A \Pi (y (\tau)) \, d\tau, \quad \Pi \in \{ \Pi_{t_0}^j, \Pi_{t_m}^j \}, \]  

(3.55)

which are related to the boundary conditions. To compute the desired bounds, we first analyze the two series in (3.52).
Lemma 3.5.15. Suppose \( \theta \in \text{int} \mathbb{B}_{\nu_u} \setminus \{0\} \) and let \( \zeta \in \mathbb{C}^{n_u} \) be such that \( \|\zeta\|_{\infty} \leq 1 \). Define a linear map \( \varphi_{u,1} : W^1_{\nu_u} \to \mathbb{C} \) by

\[
\varphi_{u,1}(p) := \sum_{l=1}^{n_u} \sum_{k \in \mathbb{N}^+_0} k_l p_k \theta^{k-e_l} \zeta_l.
\]

Then \( \varphi_{u,1} \in (W^1_{\nu_u})^* \) and

\[
\|\varphi_{u,1}\|_{\mathcal{B}(W^1_{\nu_u}, \mathbb{C})} \leq \Phi_{u,1}(\theta) := \begin{cases} 
\nu_u - 1, & \log \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right) \leq -1, \\
-\left( e \|\theta\|_{\infty} \log \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right) \right)^{-1}, & \text{otherwise.}
\end{cases}
\]

Proof. It is clear that \( \varphi_{u,1} \) is bounded, since

\[
|p_k \theta^{k-e_l}| \leq \frac{\|\theta\|_{\infty}^{-1}}{\nu_u^{|k|}} \|p\|_{\nu_u}, \quad k \in \mathbb{N}^+_0,
\]

for any \( \theta \in \text{int} \mathbb{B}_{\nu_u} \). We will use this observation to compute a bound for the operator norm. To be more precise, let \( j \in \mathbb{N}^+_0 \) be arbitrary, then

\[
|\varphi_{u,1}(\xi_j)| \leq \sum_{l=1}^{n_u} j_l |\theta|^{j-e_l} \nu_u^{-|j|} \leq \|\theta\|_{\infty}^{-1} |j| \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right)^{|j|},
\]

where \( (\xi_l)_{l \in \mathbb{N}^+_0} \) are the corner points introduced in Definition 3.2.11. In particular, observe that \( \|\varphi_{u,1}\|_{\mathcal{B}(W^1_{\nu_u}, \mathbb{C})} = \sup_{j \in \mathbb{N}^+_0 \setminus \{0\}} |\varphi_{u,1}(\xi_j)| \) by Proposition 3.2.12, since \( \varphi_{u,1}(\xi_0) = 0 \). Next, note that the map \( x \mapsto x \rho^x \), where \( \rho \in (0,1) \), is strictly increasing on \( \left[ 0, \frac{-1}{\log \rho} \right] \), strictly decreasing on \( \left[ \frac{-1}{\log \rho}, \infty \right) \) and has a global maximum on \( [0, \infty) \) at \( x = \frac{-1}{\log \rho} \). Therefore, since \( |j| \geq 1 \) for \( j \in \mathbb{N}^+_0 \setminus \{0\} \) and \( \|\theta\|_{\infty} < \nu_u \), it follows that \( \sup_{j \in \mathbb{N}^+_0 \setminus \{0\}} |\varphi_{u,1}(\xi_j)| \leq \Phi_{u,1}(\theta) \), which proves the result. \( \square \)

Remark 3.5.16. The analogs of \( \varphi_{u,1} \) and \( \Phi_{u,1} \) in the context of the stable manifold are defined similarly and are denoted by \( \varphi_{s,1} \) and \( \Phi_{s,1}(\phi) \), respectively.

Remark 3.5.17. In practice, \( \theta \) is an interval enclosure of \( \hat{\theta} \) (see Lemma 3.5.21). Recall that \( \hat{\theta} \) corresponds to a (numerically obtained) coordinate on the chart of the unstable manifold at which the to be validated piece of the connecting orbit starts. This is why we assume that \( \theta \neq 0 \).

The latter result provides a way to bound the first term in (3.52). To bound the second term (3.52), we perform a similar analysis.
Lemma 3.5.18. Suppose $\theta \in \text{int} \mathbb{B}_{\nu_u} \setminus \{0\}$ and define a linear map $\varphi_{u,2} : W^1_{\nu_u} \to \mathbb{C}$ by

$$\varphi_{u,2}(p) := \sum_{i,l=1}^{n_u} k_l (k_l - \delta_{il}) p_k \theta^{k-e_l-e_i} \tilde{\theta}_l \hat{\theta}_i.$$ 

Then $\varphi_{u,2} \in (W^1_{\nu_u})^*$ and $\|\varphi_{u,2}\|_{B(W^1_{\nu_u},\mathbb{C})} \leq \Phi_{u,2}(\theta)$, where

$$\Phi_{u,2}(\theta) := (K^u_{\text{min}} + 1)^2 \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right)^{K^u_{\text{min}} + 1}, \quad K^u_{\text{min}} + 1 \geq -2 \log \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right)^{-1},$$

$$4 \left( e \log \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right) \right)^{-2}, \quad \text{otherwise},$$

and $K^u_{\text{min}} := \min_{1 \leq i \leq n_u} K_i^u$.

Proof. The boundedness of $\varphi_{u,2}$ follows directly from the observation in (3.56) and the assumption that $\theta \in \text{int} \mathbb{B}_{\nu_u}$. To compute a bound for the operator norm, first observe that $\varphi_{u,2}(\xi_j) = 0$ for $j \in K^u_c$. Hence $\|\varphi_{u,2}\|_{B(W^1_{\nu_u},\mathbb{C})} = \sup_{j \in K^u_c} |\varphi_{u,2}(\xi_j)|$ by Proposition 3.2.12. Furthermore, a straightforward computation shows that

$$|\varphi_{u,2}(\xi_j)| \leq \sum_{i,l=1}^{n_u} j_l (j_l - \delta_{il}) |\theta|^{j-e_l-e_i} \nu_u^{-|j|},$$

$$\leq \|\theta\|^{-2}_{\infty} \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right)^{|j|} \sum_{i,l=1}^{n_u} j_l (j_l - \delta_{il})$$

$$\leq \|\theta\|^{-2}_{\infty} |j|^2 \left( \frac{\|\theta\|_{\infty}}{\nu_u} \right)^{|j|},$$

for any $j \in K^u_c$, where we used that $\left\| \hat{\theta} \right\|_{\infty}, \left\| \tilde{\theta} \right\|_{\infty} \leq 1$. Now, note that the map $x \mapsto x^2 \rho^x$, where $\rho \in (0,1)$, is strictly increasing on $[0, \frac{-2}{\log \rho}]$, strictly decreasing on $\left[ \frac{-2}{\log \rho}, \infty \right)$ and has a global maximum on $[0, \infty)$ at $x = \frac{-2}{\log \rho}$. Therefore, since $|j| \geq K^u_{\text{min}} + 1$ for $j \in K^u_c$, it follows that $\sup_{j \in K^u_c} |\varphi_{u,2}(\xi_j)| \leq \Phi_{u,2}(\theta)$, which proves the result.

Remark 3.5.19. As before, the analogs of $\varphi_{u,2}$ and $\Phi_{u,2}$ in the context of the stable manifold are denoted by $\varphi_{s,2}$ and $\Phi_{s,2}(\phi)$, respectively.
We are now ready to compute bounds for \((3.55)\). The computation of these bounds consists of a mixture of interval analysis on the computer and ordinary estimates derived with “pen and paper”. More precisely, as mentioned at the beginning of this paper, in order to manage the rounding errors on the computer, all the bounds in this paper are computed with *interval arithmetic*. Roughly speaking, this means that all the elementary operations on floating point numbers are replaced by operations on intervals with endpoints representable on a computer. In this way, one can compute rigorous bounds (and hence verify inequalities) with the aid of a computer.

Now, in order to compute bounds for \((3.55)\), we first define interval enclosures for \(\hat{\theta}, \hat{p}, \hat{\phi}\) and \(\hat{q}\). Let \(r^* > 0\) be an upper bound for the radius \(r\) and set

\[
\hat{\theta} := \prod_{j=1}^{n_u} \left[ \hat{\theta}_j - r^*, \hat{\theta}_j + r^* \right], \quad \hat{p}_k := \prod_{j=1}^{n} \left[ [\hat{p}_k]_j - \frac{r^*}{\nu_u |k|}, [\hat{p}_k]_j + \frac{r^*}{\nu_u |k|} \right], \quad k \in K^u,
\]

\[
\hat{\phi} := \prod_{j=1}^{n_s} \left[ \hat{\phi}_j - r^*, \hat{\phi}_j + r^* \right], \quad \hat{q}_k := \prod_{j=1}^{n} \left[ [\hat{q}_k]_j - \frac{r^*}{\nu_s |k|}, [\hat{q}_k]_j + \frac{r^*}{\nu_s |k|} \right], \quad k \in K^s.
\]

We require that \(r^*\) is sufficiently small and that \(\nu_u, \nu_s\) are sufficiently large so that \(\hat{\theta} \subset \text{int} \ B_{\nu_u}\) and \(\hat{\phi} \subset \text{int} \ B_{\nu_s}\).

**Remark 3.5.20.** Strictly speaking, the endpoints of the above intervals should be floating point numbers so that we can perform rigorous computations on a computer. In practice, this amounts to computing slightly larger interval enclosures for \(\hat{\theta}, \hat{p}, \hat{\phi}\) and \(\hat{q}\) (compared to the ones above). To avoid clutter in the notation, however, we have chosen to ignore this rather technical (but easily solved) issue.

**Lemma 3.5.21.** Let \(1 \leq \hat{j} \leq n\), \(0 < r \leq r^*\) and \(\Pi \in \mathbb{P}\), then

\[
\sup_{\tau \in [0,1]} \left\| \Pi A \Pi_{t_0}^\hat{j} (y(\tau)) \right\|_{\Pi(\chi_\nu)} \leq \left\| \Pi A N K \Pi_{t_0}^\hat{j} \right\|_{\Pi(\chi_\nu)} \sup_{\Pi} \beta_{u}^\hat{j}, \quad (3.57)
\]

\[
\sup_{\tau \in [0,1]} \left\| \Pi A \Pi_{m}^\hat{j} (y(\tau)) \right\|_{\Pi(\chi_\nu)} \leq \left\| \Pi A N K \Pi_{m}^\hat{j} \right\|_{\Pi(\chi_\nu)} \sup_{\Pi} \beta_{s}^\hat{j}, \quad (3.58)
\]

where

\[
\beta_{u}^\hat{j} := 2\Phi_{u,1} \left( \hat{\theta} \right) + r^* \Phi_{u,2} \left( \hat{\theta} \right) + \sum_{i,l=1}^{n_u} \sum_{k \in K^u} k_l (k_i - \delta_{il}) \left[ [\hat{p}_k]_j \right] \left[ \hat{\theta}^k \right]^{k - \epsilon_i - \epsilon_i},
\]

\[
\beta_{s}^\hat{j} := 2\Phi_{s,1} \left( \hat{\phi} \right) + r^* \Phi_{s,2} \left( \hat{\phi} \right) + \sum_{i,l=1}^{n_s} \sum_{k \in K^s} k_l (k_i - \delta_{il}) \left[ [\hat{q}_k]_j \right] \left[ \hat{\phi}^k \right]^{k - \epsilon_i - \epsilon_i}.
\]

**Proof.** Let \(\tau \in [0,1]\) and \(1 \leq \hat{j} \leq n\) be arbitrary and use Lemma 3.5.15 to see that the first term in \((3.52)\) is bounded by \(2\Phi_{u,1} \left( \hat{\theta} + \tau \hat{r} \hat{\theta} \right)\). To bound the second term
in (3.52), we first split it up into two series; one over $K_u$ and one over $K_c$. We then use Lemma 3.5.18 and the fact that $\hat{p}_k = 0$ for $k \in K_c$ to estimate

$$\left| \sum_{i,l=1, k \in \mathbb{N}_0^n} n_{il} k_l (k_i - \delta_{il}) [\hat{p}_k + \tau r \hat{p}_k]_j \left( \hat{\theta} + \tau r \hat{\theta} \right)^{k - e_i - e_i} \hat{\theta}_i \hat{\theta}_i \right|$$

$$\leq \left| \sum_{i,l=1, k \in K_u} n_{il} k_l (k_i - \delta_{il}) [\hat{p}_k + \tau r \hat{p}_k]_j \left( \hat{\theta} + \tau r \hat{\theta} \right)^{k - e_i - e_i} \hat{\theta}_i \hat{\theta}_i \right|$$

$$+ r \Phi_{u,2} \left( \hat{\theta} + \tau r \hat{\theta} \right).$$

Altogether, we conclude that

$$\left| \Pi_{t_0}^j (y(\tau)) \right| \leq 2 \Phi_{u,1} \left( \hat{\theta} + \tau r \hat{\theta} \right) + r^* \Phi_{u,2} \left( \hat{\theta} + \tau r \hat{\theta} \right)$$

$$+ \sum_{i,l=1, k \in K_u} n_{il} k_l (k_i - \delta_{il}) [\hat{p}_k + \tau r \hat{p}_k]_j \left| \hat{\theta} + \tau r \hat{\theta} \right|^{k - e_i - e_i}, \quad (3.59)$$

where we used the assumption that $r \leq r^*$. Finally, observe that $\hat{\theta} + \tau r \hat{\theta} \in \hat{\theta}$ and $\hat{p}_k + \tau r \hat{p}_k \in \hat{p}_k$, since $\left\| \hat{\theta} \right\|_\infty$, $\left\| \hat{p} \right\|_{W^1_{u,n}} \leq 1$. Hence (3.59) is contained in $\beta^*_{u}$ for all $\tau \in [0,1]$, which proves the result. \qed

**Remark 3.5.22.** Observe that the computation of the bound in this lemma is finite, since $K_u$ is a finite set of multi-indices.

**Chebyshev coefficients** Next, we consider the terms associated to

$$\int_0^1 \Pi_{a}^{ij} (y(\tau)) \, d\tau, \quad 1 \leq i \leq m, \ 1 \leq j \leq n, \quad (3.60)$$

which are related to the equations for the Chebyshev coefficients. We start by computing a bound for (3.53). Since we will have to perform a similar analysis for the Taylor coefficients in the next paragraph, we first state a general result.

**Lemma 3.5.23.** Suppose $(X, *)$ is a Banach algebra. Let $1 \leq j \leq n$ and define $G_j : \bigoplus_{l=1}^n X \rightarrow X$ and $\tilde{g}_j : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$G_j(x) := \sum_{\alpha \in A} [g_\alpha]_j x^\alpha, \quad \tilde{g}_j(x) := \sum_{\alpha \in A} \left| [g_\alpha]_j \right| x^\alpha,$$

where $\{g_\alpha : \alpha \in A\} \subset \mathbb{R}^n$ are the coefficients of $g$ in the monomial basis. Then

$$\left\| D^2 G_j (x + \tau rz) \right\| \leq D^2 \tilde{g}_j \left( \left\| x_1 \right\|_X + r^*, \ldots, \left\| x_n \right\|_X + r^* \right) \left[ 1_n, 1_n \right],$$

for any $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n$, $z = (z_i)_{i=1}^n \in \bigoplus_{l=1}^n X$ such that $\left\| y \right\|, \left\| z \right\| \leq 1$, $\tau \in [0,1]$ and $0 < r \leq r^*$. Here $\left\| \cdot \right\|$ denotes the max-norm on $\bigoplus_{l=1}^n X$. \hfill 122
Therefore, since

\[ D^2 G_j (x + \tau rz) [y, z] = \sum_{i,l=1}^{n} \sum_{\alpha \in \mathcal{A}} a_i (x_i - \delta_{il}) [g_{\alpha}]_j (x + \tau rz)^{\alpha - e_i - e_l} y_i * z_l. \]

Hence, by the Banach algebra estimate,

\[ \| D^2 G_j (x + \tau rz) [y, z] \| \leq \sum_{i,l=1}^{n} \sum_{\alpha \in \mathcal{A}} a_i (x_i - \delta_{il}) \| [g_{\alpha}]_j \| \prod_{k=1}^{n} (\| x_k \|_X + r^*)^{(\alpha - e_i - e_l)_k} \]

\[ = D^2 \tilde{g}_j (\| x_1 \|_X + r^*, \ldots, \| x_n \|_X + r^*) [1_n, 1_n], \]

since \( \| y \|, \| z \| \leq 1, \tau \in [0,1] \) and \( 0 < r \leq r^* \).

**Remark 3.5.24.** Note that the convolution mappings \([C]_j\) and \([c]_j\) are of the form \( G_j \).

Next, we use the above result to compute bounds for (3.60). The key observation is stated in the next lemma.

**Lemma 3.5.25.** Let \( 1 \leq \bar{i} \leq m \), \( 1 \leq \bar{j} \leq n \) and \( 0 < r \leq r^* \), then

\[
\sup_{\tau \in [0,1]} \| \Pi_{\tilde{a}_\tau}^{\mathbb{T}_N} (y(\tau)) \|_{\nu_{t_i}} \leq \frac{L\nu_t (t_{\bar{i}} - t_{\bar{i}-1})}{2} D^2 \tilde{g}_j \left( \| a^{\bar{i}}_1 \|_{\nu_{t_i}} + r^*, \ldots, \| a^{\bar{i}}_n \|_{\nu_{t_i}} + r^* \right) [1_n, 1_n].
\]

**Proof.** Define linear operators \( \sigma_1 : \ell^1_{\nu_{t_i}} \to \ell^1_{\nu_{t_i}} / \mathbb{C} \) and \( \{ \sigma_{\bar{i}} : \ell^1_{\nu_{t_i}} \to \ell^1_{\nu_{t_i}} \}_{\bar{i}=2}^{m} \) by

\[
\sigma_1 (a^1) := (a^1_{k-1} - a^1_{k+1})_{k \in \mathbb{N}}, \quad \sigma_{\bar{i}} (a^{\bar{i}}) := \begin{bmatrix} 0 \\ (a^{\bar{i}}_{k-1} - a^{\bar{i}}_{k+1})_{k \in \mathbb{N}} \end{bmatrix}.
\]

A direct application of Proposition 3.2.9 shows that these operators are bounded and that

\[
\| \sigma_1 \|_{\mathcal{B}(\ell^1_{\nu_{t_i}}, \ell^1_{\nu_{t_i}} / \mathbb{C})} = 2\nu_{t_i}, \quad \| \sigma_{\bar{i}} \|_{\mathcal{B}(\ell^1_{\nu_{t_i}}, \ell^1_{\nu_{t_i}})} = 2\nu_{t_i}, \quad 2 \leq \bar{i} \leq m. \tag{3.61}
\]

Therefore, since

\[
\Pi_{\tilde{a}_\tau}^{\mathbb{T}_N} (y(\tau)) = \frac{L(t_{\bar{i}} - t_{\bar{i}-1})}{4} \sigma_{\bar{i}} D^2 [c]_{\bar{j}} (\tilde{a}^{\bar{i}} + \tau r \tilde{a}^{\bar{i}}) [\tilde{a}^{\bar{i}}, \tilde{a}^{\bar{i}}]
\]

by (3.53), the result follows directly from (3.61) and Lemma 3.5.23.

It is now a straightforward task to compute bounds for (3.60).
Lemma 3.5.26 (Scalar and Taylor projections). Let $1 \leq i \leq m$, $1 \leq j \leq n$, $0 < r \leq r^*$ and $\Pi \in \left\{ \Pi^j_\ell^i, \Pi^i_\ell^j, \Pi^{i,j}_\ell^i_\ell^j, \Pi^{i,j}_{\lambda,\lambda}, \Pi^{i,j}_{\lambda,\rho}, \Pi^{i,j}_{\rho,\rho} \right\}$, then

$$\sup_{\tau \in [0,1]} \left\| \Pi A \Pi^{i,j}_a (y(\tau)) \right\|_{\Pi(\mathcal{X}_\nu)} \leq \left\| \Pi A_{NK} \Pi^{i,j}_{[1:N_i-1]} \right\|_{B(\epsilon_{\nu_i}^i,\Pi(\mathcal{X}_\nu))}$$

$$+ \frac{L\nu_i (t_i - t_{i-1})}{2} D^2 \tilde{g}_j \left( \left\| \begin{bmatrix} \tilde{\alpha} \end{bmatrix} \right\|_{\nu_i} + r^* \right) \left\| \begin{bmatrix} \tilde{\alpha} \end{bmatrix} \right\|_{\nu_i} + r^*) \begin{bmatrix} 1_n, 1_n \end{bmatrix}.$$  

Proof. It suffices to observe that

$$\Pi A \Pi^{i,j}_a (y(\tau)) = \Pi A_{NK} \Pi^{i,j}_{[1:N_i-1]} (y(\tau)), \quad \Pi \in \left\{ \Pi^j_\ell^i, \Pi^i_\ell^j, \Pi^{i,j}_\ell^i_\ell^j, \Pi^{i,j}_{\lambda,\lambda}, \Pi^{i,j}_{\lambda,\rho}, \Pi^{i,j}_{\rho,\rho} \right\},$$

by construction of the approximate inverse $A$. Therefore, the result follows directly from Lemma 3.5.23.

Lemma 3.5.27 (Chebyshev projections). Let $1 \leq i, i \leq m$, $1 \leq j, j \leq n$ and $0 < r \leq r^*$, then

$$\sup_{\tau \in [0,1]} \left\| \Pi^j_a A \Pi^{i,j}_a (y(\tau)) \right\|_{\nu_i}$$

$$\leq \frac{L\nu_i (t_i - t_{i-1})}{2} D^2 \tilde{g}_j \left( \left\| \begin{bmatrix} \tilde{\alpha} \end{bmatrix} \right\|_{\nu_i} + r^* \right) \left\| \begin{bmatrix} \tilde{\alpha} \end{bmatrix} \right\|_{\nu_i} + r^*) \begin{bmatrix} 1_n, 1_n \end{bmatrix}.$$

$$\cdot \max \left\{ \left\| \Pi^j_a A_{NK} \Pi^{i,j}_{[1:N_i-1]} \right\|_{B(\epsilon_{\nu_i}^i,\Pi(\mathcal{X}_\nu))} \begin{bmatrix} 1_n, 1_n \end{bmatrix}, \begin{bmatrix} 1_{N_i}, 1_{N_i} \end{bmatrix}, \begin{bmatrix} 1_{N_i+1}, 1_{N_i+1} \end{bmatrix} \right\}, \quad (\bar{i}, \bar{j}) = (i, j).$$

Proof. It follows from the definition of the approximate inverse that

$$\Pi^j_a A \Pi^{i,j}_a = \Pi^j_a A_{NK} \Pi^{i,j}_{[1:N_i-1]} + \Pi^j_a A \Pi^{i,j}_{[N_i: \infty]}$$

$$= \begin{cases} \Pi^j_a A_{NK} \Pi^{i,j}_{[1:N_i-1]}, & (\bar{i}, \bar{j}) \neq (i, j), \\ \Pi^j_a A_{NK} \Pi^{i,j}_{[1:N_i-1]} & (i, j) = (i, j) \end{cases}$$

$$\begin{cases} \frac{1}{N_i} & (\bar{i}, \bar{j}) \neq (i, j), \\ \frac{1}{N_i + 1} & (i, j) = (i, j). \end{cases}$$

124
Lemma 3.5.28. Let \( \Pi \) a manifold. Observe that \( \Pi \) which are related to the equations for the Taylor coefficients of the (un)stable of (3.62), we introduce the term \( \Pi \), \( y \)

\[
\| \Pi_{a}^{ij} A \Pi_{a}^{j\ell} \|_{B(\ell_{v}_{q}, \ell_{v}_{s})} = \left\{ \begin{array}{ll}
\| \Pi_{a}^{ij} A_{NK} \Pi_{a}^{j[1:N_{t}-1]} \|_{B(\ell_{v}_{q}, \ell_{v}_{s})}, & (i, j) \neq (i, j), \\
\max \left\{ \| \Pi_{a}^{ij} A_{NK} \Pi_{a}^{j[1:N_{t}-1]} \|_{B(\ell_{v}_{q}, \ell_{v}_{s})}, \frac{1}{N_{t}} \right\}, & (i, j) = (i, j),
\end{array} \right.
\]

by Proposition 3.2.9. The result now follows from Lemma 3.5.23. \( \square \)

**Taylor coefficients** Finally, we consider the terms associated to

\[
\int_{0}^{1} A \Pi (y (\tau)) d \tau, \quad \Pi \in \left\{ \Pi_{P}^{j}, \Pi_{Q}^{j} \right\},
\]

which are related to the equations for the Taylor coefficients of the (un)stable manifolds. Observe that \( \Pi_{P}^{j} (y (\tau)) \notin W_{\nu_{p}}^{1} \), due to the presence of the terms \( \langle \hat{\lambda}^{u}, k \rangle \hat{p}_{k} + \langle \hat{\lambda}^{u}, k \rangle \hat{p}_{k} \), see (3.54). For this reason, in order to facilitate the analysis of (3.62), we introduce the term \( y_{P}^{j} (\tau) := \Pi_{P}^{j} (y (\tau)) - y^{j}_{P, \lambda} (\tau) \), where

\[
\left( y_{P, \lambda}^{j} (\tau) \right)_{k} := \begin{cases} 0, & k \in K^{u}, \\
\langle \hat{\lambda}^{u}, k \rangle [\hat{p}_{k}]_{j} + \langle \hat{\lambda}^{u}, k \rangle [\hat{p}_{k}]_{j}, & k \notin K^{u},
\end{cases}
\]

i.e., \( y_{P}^{j} (\tau) \) is defined by removing the linear terms \( \langle \hat{\lambda}^{u}, k \rangle [\hat{p}_{k}]_{j} + \langle \hat{\lambda}^{u}, k \rangle [\hat{p}_{k}]_{j} \) from the tail of \( \Pi_{P}^{j} (y (\tau)) \). Hence \( y_{P}^{j} (\tau) \in W_{\nu_{p}}^{1} \). An analogous decomposition \( \Pi_{Q}^{j} (y (\tau)) = y_{Q}^{j} (\tau) + y_{Q, \lambda}^{j} (\tau) \) is used to analyze the terms associated to the stable manifold.

Lemma 3.5.28. Let \( 1 \leq j \leq n \) and \( 0 < r \leq r^{*} \), then

\[
\sup_{\tau \in [0, 1]} \left\| y_{P}^{j} (\tau) \right\|_{\nu_{u}} \leq \sup_{\tau \in [0, 1]} \left\| \sigma_{u}^{j} \right\|_{\nu_{u}}, \quad \sup_{\tau \in [0, 1]} \left\| y_{Q}^{j} (\tau) \right\|_{\nu_{u}} \leq \sup_{\tau \in [0, 1]} \left\| \sigma_{s}^{j} \right\|_{\nu_{s}},
\]

where

\[
\sigma_{u}^{j} := \sum_{i_{3}=1}^{n} \sum_{i_{1}=1}^{n_{u}} \left| [\hat{p}_{c_{i_{3}}}^{j}]_{i_{3}} \nu_{u} + r^{*} \right| \sum_{i_{1}, i_{2}=1}^{n} \left| \frac{\partial^{3} g_{j}}{\partial x_{i_{1}} \partial x_{i_{2}} \partial x_{i_{3}}} (\hat{p}_{0}) \right| + 3 \sum_{i_{1}, i_{2}=1}^{n} \left| \frac{\partial^{2} g_{j}}{\partial x_{i_{1}} \partial x_{i_{2}}} (\hat{p}_{0}) \right| + 2 (|K^{u}| + 1)
\]

\[
+ D^{2} g_{j} \left( \left\| [\hat{p}]_{1} \right\|_{\nu_{u}} + r^{*}, \ldots, \left\| [\hat{p}]_{n} \right\|_{\nu_{u}} + r^{*} \right) \left[ 1_{n}, 1_{n} \right],
\]

125
The bounds $\{\sigma_s^j : 1 \leq j \leq n\}$ associated to the stable manifold are defined analogously.

**Proof.** Let $\tau \in [0, 1]$ and $1 \leq j \leq n$ be arbitrary. It follows directly from (3.54) that

$$\left| \left( y^j_P (\tau) \right)_0 \right| \leq \sum_{i_1, i_2 = 1}^{n} \left| \frac{\partial^2 g_j}{\partial x_{i_1} \partial x_{i_2}} (\hat{p}_0 + \tau \hat{r} \hat{p}_0) \right|,$$

(3.63)

since $\|\hat{p}\|_{\nu_u,n}, \|\hat{p}\|_{\nu_u,n} \leq 1$. Next, we consider the first order components of $y^j_P(\tau)$ by separately analyzing the terms in (3.54) for $|k| = 1$. To this end, observe that

$$\sum_{i_1,i_2=1}^{n} \left| \sum_{i=1}^{n} \left( \sum_{i_3=1}^{n} \left[ \hat{p}_{e_i} \right]_{i_3} \right) \nu_u + r^* \right| \sum_{i_1,i_2=1}^{n} \left| \frac{\partial^3 g_j}{\partial x_{i_1} \partial x_{i_2} \partial x_{i_3}} (\hat{p}_0 + \tau \hat{r} \hat{p}_0) \right|,$$

and

$$\sum_{i_1,i_2=1}^{n} \left| \sum_{i_3=1}^{n} \left( \sum_{i=1}^{n} \left[ \hat{p}_{e_i} \right]_{i_3} \right) \nu_u \right| \leq \sum_{i_1,i_2=1}^{n} \left| \frac{\partial^2 g_j}{\partial x_{i_1} \partial x_{i_2}} (\hat{p}_0 + \tau \hat{r} \hat{p}_0) \right|,$$

(3.64)

where we used that $r \leq r^*$, $\left\| \hat{\lambda}_i \right\|_{\infty}, \left\| \hat{\lambda}'_i \right\|_{\infty}, \|\hat{p}\|_{\nu_u,n}, \|\hat{p}\|_{\nu_u,n} \leq 1$. It now follows from the expression in (3.54) that

$$\sum_{k \in K_u^0 \atop |k| = 1} \left| \left( y^j_P (\tau) \right)_k \right| \nu_u$$

$$\leq \sum_{i_1,i_2=1}^{n} \left| \sum_{i_3=1}^{n} \left( \sum_{i=1}^{n} \left[ \hat{p}_{e_i} \right]_{i_3} \right) \nu_u + r^* \right| \sum_{i_1,i_2=1}^{n} \left| \frac{\partial^3 g_j}{\partial x_{i_1} \partial x_{i_2} \partial x_{i_3}} (\hat{p}_0 + \tau \hat{r} \hat{p}_0) \right|$$

$$+ 2 \sum_{i_1,i_2=1}^{n} \left| \frac{\partial^2 g_j}{\partial x_{i_1} \partial x_{i_2}} (\hat{p}_0 + \tau \hat{r} \hat{p}_0) \right| + 2.$$

Finally, we consider the remainder of $y^j_P$, i.e., the components associated to $|k| \geq 2$. First, note that

$$\sum_{k \in K_u^0 \atop |k| \geq 2} \left| \left\langle \hat{\lambda}_i, k \right| \right\| \nu_u k \leq 2 \|K_u^0 \|,$$
since \( \| \tilde{\lambda}^u \|_\infty, \| \lambda^u \|_\infty, \| \tilde{\lambda} \|_{\nu_{u,n}}, \| \lambda \|_{\nu_{u,n}} \leq 1 \) and \( |k| \leq |K^u| \) for \( k \in K^u \). Furthermore,

\[
\sum_{k \in \mathbb{N}^n_u |k| \geq 2} D^2 C_k (\tilde{\lambda} + \tau \tilde{\lambda}) [\dot{\lambda}^u, \dot{\lambda}^s] \leq D^2 \tilde{g}_j \left( \| \tilde{\lambda} \|_{\nu_u} + r^*, \ldots, \| \tilde{\lambda} \|_{\nu_u} + r^* \right) [1_n, 1_n]
\]

by Lemma 3.5.23. Therefore,

\[
\sum_{k \in \mathbb{N}^n_u |k| \geq 2} \left( \left| y^j_P(\tau) \right| \right)_k \nu^k_u \leq 2 |K^u| + D^2 \tilde{g}_j \left( \| \tilde{\lambda} \|_{\nu_u} + r^*, \ldots, \| \tilde{\lambda} \|_{\nu_u} + r^* \right) [1_n, 1_n].
\] (3.65)

Altogether, the sum of (3.63), (3.64) and (3.65) constitutes an upper bound for \( \left\| y^j_P(\tau) \right\|_{\nu_u} \), for any \( \tau \in [0, 1] \) and \( r \leq r^* \), and is contained in \( \sigma^j_u \). This proves the result.

It is now a straightforward task to compute bounds for (3.62).

**Lemma 3.5.29.** Let \( 1 \leq j \leq n \) and \( 0 < r \leq r^* \), then

\[
\sup_{\tau \in [0,1]} \left\| \Pi A \Pi^j_P (y(\tau)) \right\|_{\Pi(\chi_u)} \leq \left\| \Pi A_{NK} \Pi^j_P \right\|_{\mathcal{B}(W^1_{\nu_u}, \Pi(\chi_u))} \sup \sigma^j_u,
\] (3.66)

for \( \Pi \in \left\{ \Pi^j_{\theta}, \Pi^j_{\phi}, \Pi^j_{\lambda_u}, \Pi^j_{\lambda_s}, \Pi^j_{\mu_a}, \Pi^j_Q \right\} \) and

\[
\sup_{\tau \in [0,1]} \left\| \Pi A \Pi^j_Q (y(\tau)) \right\|_{\Pi(\chi_u)} \leq \left\| \Pi A_{NK} \Pi^j_Q \right\|_{\mathcal{B}(W^1_{\nu_u}, \Pi(\chi_u))} \sup \sigma^j_s
\] (3.67)

for \( \Pi \in \left\{ \Pi^j_{\theta}, \Pi^j_{\phi}, \Pi^j_{\lambda_u}, \Pi^j_{\lambda_s}, \Pi^j_{\mu_a}, \Pi^j_Q \right\} \).

**Proof.** It suffices to observe that

\[
\Pi A \Pi^j_P (y(\tau)) = \Pi A_{NK} \Pi^j_P (y(\tau)), \quad \Pi \in \left\{ \Pi^j_{\theta}, \Pi^j_{\phi}, \Pi^j_{\lambda_u}, \Pi^j_{\lambda_s}, \Pi^j_{\mu_a}, \Pi^j_Q \right\},
\]

by construction of the approximate inverse \( A \). Therefore, since

\[
\Pi^j_{NK} (y(\tau)) = \Pi^j_P \left( y^j_P(\tau) \right),
\]

the result follows from Lemma 3.5.28. \( \square \)

127
Lemma 3.5.30. Let $1 \leq j, \tilde{j} \leq n$ and $0 < r \leq r^*$, then

$$
\sup_{\tau \in [0,1]} \left\| \Pi_{\tau}^j A \Pi_{\tau}^j (y(\tau)) \right\|_{\nu_u}
\leq \left\{ \begin{array}{ll}
\left\| \Pi_{\tau}^j A N K \Pi_{\tau}^j \right\|_{B(W_1^{\nu_u}, W_1^{\nu_u})} & \text{if } \tilde{j} \neq j,
\max \left\{ \left( \min_{1 \leq i \leq n_u} \left| \Re \left( \hat{\lambda}_i^u \right) \right| \left( K_i^u + 1 \right) \right)^{-1} \right\} \sup \sigma_{\tilde{j}},
+ 2 \left( \min_{1 \leq i \leq n_u} \left| \Re \left( \hat{\lambda}_i^u \right) \right| (K_i^u + 1) \right)^{-1} & \text{if } \tilde{j} = j.
\end{array} \right.
$$

The statement and corresponding bound for the stable manifold is analogous.

Proof. If $\tilde{j} \neq j$, then

$$
\Pi_{\tau}^j A \Pi_{\tau}^j (y(\tau)) = \Pi_{\tau}^j A N K \Pi_{\tau}^j \left( y_{\tilde{j}} (\tau) \right)
$$

by definition of the approximate inverse and $y_{\tilde{j}} (\tau)$. Therefore, in this case, the result follows immediately from Lemma 3.5.28. Now, assume that $\tilde{j} = j$ and observe that

$$
\Pi_{\tau}^j A \Pi_{\tau}^j (y(\tau)) = \Pi_{\tau}^j A \Pi_{\tau}^j \left( y_{\tilde{j}} (\tau) \right) + \Pi_{\tau}^j A \Pi_{\tau}^j \left( y_{j, \lambda} (\tau) \right).
$$

In particular, since $y_{\tilde{j}} (\tau) \in W_1^{\nu_u}$, we may bound the first term in the above expression by

$$
\left\| \Pi_{\tau}^j A \Pi_{\tau}^j \left( y_{\tilde{j}} (\tau) \right) \right\|_{\nu_u} \leq \left\| \Pi_{\tau}^j A \Pi_{\tau}^j \right\|_{B(W_1^{\nu_u}, W_1^{\nu_u})} \left\| y_{\tilde{j}} (\tau) \right\|_{\nu_u}
\leq \left\| \Pi_{\tau}^j A \Pi_{\tau}^j \right\|_{B(W_1^{\nu_u}, W_1^{\nu_u})} \sup \sigma_{\tilde{j}},
$$

where in the last line we used Lemma 3.5.28 again.

Next, we derive a more explicit expression for $\left\| \Pi_{\tau}^j A \Pi_{\tau}^j \right\|_{B(W_1^{\nu_u}, W_1^{\nu_u})}$. To this end, recall the definition of $\Pi_{\tau}^j A \Pi_{\tau}^j$ (see Definition 3.4.6) and observe that

$$
\left| \left\langle \hat{\lambda}_i^u, k \right\rangle \right| \geq \left( \min_{1 \leq i \leq n_u} \left| \Re \left( \hat{\lambda}_i^u \right) \right| \left( K_i^u + 1 \right) \right)^{-1}
$$

for any $k \in K_i^u$. Therefore, it follows from Proposition 3.2.12 that

$$
\left\| \Pi_{\tau}^j A \Pi_{\tau}^j \right\|_{B(W_1^{\nu_u}, W_1^{\nu_u})} = \max \left\{ \left\| \Pi_{\tau}^j A N K \Pi_{\tau}^j \right\|_{B(W_1^{\nu_u}, W_1^{\nu_u})}, \left( \min_{1 \leq i \leq n_u} \left| \Re \left( \hat{\lambda}_i^u \right) \right| \left( K_i^u + 1 \right) \right)^{-1} \right\}.
$$
Finally, a straightforward computation shows that
\[
\left\| \Pi_P^j A \Pi_P^j \left( y_{P, \lambda}^j (\tau) \right) \right\|_{\nu_u} = \sum_{k \in K_u^\tau} \left| \langle \hat{\lambda}^u, k \rangle [\hat{p}_k]^j + \langle \hat{\lambda}^u, k \rangle [\tilde{p}_k]^j \right| \left| \langle \hat{\lambda}^u, k \rangle \right|^{-1} \nu_u^k
\]
\[
\leq 2 \left( \min_{1 \leq i \leq n_u} \left| \Re \left( \hat{\lambda}^u_i \right) \right| \right)^{-1},
\]
since \(|\langle \lambda^u, k \rangle| \leq \|\lambda^u\|_{\infty} |k|\) and \(\left| \langle \hat{\lambda}^u, k \rangle \right| \geq \min_{1 \leq i \leq n_u} \left| \Re \left( \hat{\lambda}^u_i \right) \right| |k|\) for any \(k \in \mathbb{N}_0^{n_u}\) and \(\lambda^u \in \mathbb{C}^{n_u}\). Altogether, this proves the result.

**Second order coefficients of \(Z_\Pi(r)\)**

We are now ready to finish the construction of the quadratic polynomials \(Z_\Pi(r)\) for \(\Pi \in \mathbb{P}\). As before, we first introduce some additional notation. We will denote the bounds in (3.57), (3.58), (3.66), (3.67) and Lemmas 3.5.26, 3.5.27, 3.5.30 by \(Z^{2, \tilde{\jmath} t_0}_{\Pi}, Z^{2, \tilde{\jmath} t_m}_{\Pi}, Z^{2, \tilde{\jmath} j_P}_{\Pi}, Z^{2, \tilde{\jmath} Q}_{\Pi}, Z^{2, \iota \tilde{\jmath} N}_{\Pi}, Z^{2, \iota \tilde{\jmath} N}_{\Pi} \) and \(Z^{2, \tilde{\jmath} P}_{\Pi}, Z^{2, \tilde{\jmath} Q}_{\Pi}\) respectively. Finally, we set
\[
Z^2_{\Pi} := \sum_{j=1}^n \left( Z^{2, \tilde{\jmath} t_0}_{\Pi} + Z^{2, \tilde{\jmath} t_m}_{\Pi} + \sum_{i=1}^m Z^{2, \iota \tilde{\jmath} N}_{\Pi} + Z^{2, \tilde{\jmath} P}_{\Pi} + Z^{2, \tilde{\jmath} Q}_{\Pi} \right), \quad \Pi \in \mathbb{P},
\]
and define
\[
Z_\Pi(r) := Z^1_{\Pi} + Z^2_{\Pi} r, \quad \Pi \in \mathbb{P}.
\]
Then \(Z_\Pi(r)\) satisfies (3.27) by construction.

### 3.6 Applications: traveling fronts in parabolic PDEs

In this section we use our method to prove the existence of connecting orbits in systems of ODEs which arise from the study of traveling fronts in scalar parabolic PDEs. In addition, we perform (non-rigorous) continuation and demonstrate the effectiveness of the phase condition introduced in Definition 3.3.14. Before we proceed to the applications, we first give a rough outline of our main procedure for validating connecting orbits. We have tried to automate as many steps as possible, but there are still certain steps which are based on experimentation.

**Step 1: Compute parameterizations of the local (un)stable manifolds**

1. Compute numerical approximations \(\bar{p}_0\) and \(\bar{q}_0\) of the equilibria of interest.

2. Compute numerical approximations
\[
\left\{ \left( \tilde{\lambda}^u_k, \bar{p}_k \right) : |k| = 1 \right\}, \quad \left\{ \left( \tilde{\lambda}^u_k, \bar{q}_k \right) : |k| = 1 \right\},
\]
of the eigendata associated to $Dg(p_0)$ and $Dg(q_0)$, respectively. In this step we set the length of the approximate eigenvectors to one.

1.3 Choose the number of Taylor coefficients $K^u \in \mathbb{N}^{n_u}$ and $K^s \in \mathbb{N}^{n_s}$ and compute approximate zeros $\left(\hat{\lambda}^u, \hat{p}\right)$, $\left(\hat{\lambda}^s, \hat{q}\right)$ of the mappings

\[
(\lambda^u, p) \mapsto \left[\Pi^{jK_u} F_P (\lambda^u, p)\right]_{j=1}^{n_u}, \quad (\lambda^s, q) \mapsto \left[\Pi^{jK_s} F_Q (\lambda^s, q)\right]_{j=1}^{n_s}
\]

by Newton’s method.

1.4 If necessary, increase the truncation parameters and rescale the eigenvectors so that validation is feasible, see Remark 3.6.1 below.

**Step 2: Compute an accurate approximation of a connecting orbit**

2.1 Compute a numerical approximation of a connecting orbit. This step is based on solving the truly nonlinear part of the problem and involves experimentation. It is obviously problem dependent.

2.2 Use the domain decomposition algorithm developed in Chapter 2 to compute a grid $(t_i)_{i=0}^m$ and an accurate approximate connecting orbit

\[
\hat{u} = \sum_{i=1}^m 1_{[t_{i-1}, t_i]} \left(\hat{a}_0 + 2 \sum_{k=1}^{N_i-1} \hat{a}_k T_k^i\right),
\]

so that the decay rates of the Chebyshev coefficients $\hat{a}_i$ are equidistributed over the subdomains $[t_{i-1}, t_i]$. The number of modes $N_i$ is chosen in such a way that $|\hat{a}_{N_i-1}| \approx 10^{-16}$. The number of subdomains $m$ is determined by experimentation. In general, we use as many subdomains as necessary in order to ensure high decay rates of the Chebyshev coefficients.

2.3 Use $\hat{u}$ as a reference orbit to fix the time parameterization of the connecting orbit (see Definition 3.3.14).

**Step 3: Validate the connecting orbit and (un)stable manifolds**

3.1 Combine the results from the previous two steps to construct a symmetric approximate zero $\hat{x} = (\hat{\theta}, \hat{\phi}, \hat{\lambda}^u, \hat{\lambda}^s, \hat{a}, \hat{p}, \hat{q})$ of $F_{NK}$ (see Remark 3.3.20).

3.2 Set $r^* = 10^{-5}$, $\nu_u = \nu_s = 1$ (see Remark 3.6.1) and compute the weights $(\nu_i)_{i=1}^m$ as explained in Remark 3.6.2 below.

3.4 Initialize the numerical data with interval arithmetic and construct the radii-polynomials

\[
p_{\Pi}(r) := Z_{\Pi}^2 r^2 + (Z_{\Pi}^1 - 1) r + Y_{\Pi}, \quad \Pi \in \mathbb{P}.
\]
3.5 Determine an interval $I$ on which all the radii polynomials are negative.

If we fail to find an interval $I$ on which all the radii polynomials are negative, we try to determine which parameters (the truncation parameters, the weights $\nu_i$ or the “scalings” of the coefficients $\hat{p}$ and $\hat{q}$) need to be modified by “visual” inspection and try again.

Remark 3.6.1 (Scaling and the number of Taylor coefficients). Observe that for any $\nu_u, \tilde{\nu}_u > 0$ it holds that $p \in W^1_{\nu_u}$ if and only if $\frac{\nu_u}{\tilde{\nu}_u} p \in W^1_{\tilde{\nu}_u}$ with the scaling notation introduced in Remark 3.3.3. Therefore, since the parameterization mappings $F_P$ and $F_Q$ are invariant under the rescaling $p \mapsto \mu p$ (see Remark 3.3.3), we have chosen to set $\nu_u = \nu_s = 1$ and search for appropriate scalings which ensure that $\hat{p}$ and $\hat{q}$ decay sufficiently fast to zero. To be more precise, we explain in detail how we choose the scalings of the eigenvectors and the number of Taylor coefficients for the unstable manifold (the procedure for the stable manifold is analogous).

The main idea is to choose the scalings and number of Taylor coefficients in such a way that the bound for $\Pi_f^j \left( DF(\hat{z}) - \hat{D}F \right)$ is below some prescribed tolerance. More precisely, in light of (3.38), (3.48) and Lemma 3.5.12, we aim to find a truncation parameter $K_u^u \in \mathbb{N}_0^{n_u}$ and a scaling factor $\mu \in (0, \infty)^{n_u}$ such that

$$
\left[ \min_{1 \leq i \leq n_u} (K_i^u + 1) \min_{1 \leq i \leq n_u} \left| \text{Re} \left( \lambda_i \right) \right| \right]^{-1} \left\| \mu \hat{G}^{jl} \right\|_1 \leq \epsilon_u, \quad 1 \leq j, l \leq n, \tag{3.68}
$$

where $\epsilon_u > 0$ (in practice we set $\epsilon_u = \frac{1}{2}$). We start by determining $K_u^u \in \mathbb{N}_0^{n_u}$. To this end, observe that the scaling factor $\mu$ has no effect on $\left| \hat{G}^{jl}_0 \right|$. For this reason, we set $K_u^u := \left( \max_{1 \leq j, l \leq n_u} K_{jl}^u \right) 1_{n_u}$, where $K_{jl}^u \in \mathbb{N}$ is the smallest integer such that

$$
K_{jl}^u > \frac{\left| \hat{G}^{jl}_0 \right|}{\zeta \epsilon_u \min_{1 \leq i \leq n_u} \left| \text{Re} \left( \lambda_i \right) \right|} - 1, \quad 1 \leq j, l \leq n, \quad \zeta \in (0, 1].
$$

Here $\zeta \in (0, 1]$ is an additional parameter chosen through experimentation (in practice we use $\zeta = \frac{3}{4}$).

Next, we determine an appropriate scaling factor $\mu$. Let $1 \leq j, l \leq n$ and approximate $\left| \hat{G}_k^{jl} \right| \approx \left| \hat{G}_0^{jl} \rho_{jl}^{-|k|} \right|$, where $\rho_{jl} = e^{-s_{jl}}$ and $s_{jl}$ is the slope of the best line through the points

$$
\left\{ \left( d, \log \left( \sum_{|k|=d} \left| \hat{G}_k^{jl} \right| \right) \right) : \sum_{|k|=d} \left| \hat{G}_k^{jl} \right| > 10^{-16}, \ 0 \leq d \leq |M_{jl}K_u^u| \right\}. \tag{3.69}
$$

Recall that $M_{jl} = \text{order} \left( \frac{\partial g_j}{\partial x_l} \right)$. Now, if $\mu_i < \rho_{jl}$ for all $1 \leq i \leq n_u$ and $|K_u^u|$ is
sufficiently large, then
\[ \left\| \mu \hat{G}^{jl} \right\|_1 \approx \left| \hat{G}^{jl}_0 \prod_{i=1}^{n_u} \frac{\rho_{jl}}{\rho_{jl} - \mu} \right|. \]

Motivated by this observation and the inequality in (3.68), we set \( \mu_i = \mu \) for \( 1 \leq i \leq n_u \) and require that
\[ \frac{\rho_{jl}}{\rho_{jl} - \mu} \leq \left[ \frac{\varepsilon_u \min_{1 \leq i \leq n_u} (K_i^{1u} + 1) \min_{1 \leq i \leq n_u} |\text{Re} (\hat{\lambda}_l^{i})|}{|\hat{G}^{jl}_0|} \right]^{\frac{1}{2}} =: \xi_{jl} \]
for all \( 1 \leq j, l \leq n \). Therefore, we set
\[ \mu := \min_{1 \leq j, l \leq n_u} \frac{\rho_{jl} \xi_{jl} - 1}{\xi_{jl}}. \]

We remark that one could determine a more “refined” scaling factor \( \mu \), which need not be the same in each direction, by taking the decay rates of \( \hat{G}^{jl} \) in each separate direction into account (as opposed to using the “uniform” rate in (3.69) which ignores the different directions of the array). In addition, one could take the different sizes of the eigenvalues into account in the definition of \( \xi_{jl} \).

Remark 3.6.2. To determine the weights \( (\nu_i)_{i=1}^m \), we use a heuristic procedure slightly more refined than the one used in Chapter 2. Namely, we try to ensure that the bound for the tail of \( \Pi_{a}^{ij} \left( DF (\hat{x}) - DF \right) \) is below some prescribed tolerance (rather than requiring the residual to be below some tolerance as in Chapter 2). More precisely, in light of (3.37), we require that
\[ \frac{L(t_i - t_{i-1})}{2N_i} \left( \nu_i + \nu_i^{-1} \right) \| \hat{g}^{ijl} \|_{\nu_i} \leq \varepsilon, \quad 1 \leq i \leq m, \ 1 \leq j, l \leq n. \] (3.70)
where \( \varepsilon > 0 \) is some prescribed tolerance (in practice we set \( \varepsilon = \frac{1}{2} \)).

We use the rough approximation \( \| \hat{g}^{ijl} \|_{\nu_i} \approx |\hat{g}^{ijl}_0| \rho^{-k}_{ijl} \), where \( \rho_{ijl} = e^{-s_{ijl}} \) and \( s_{ijl} \) is the slope of the best line through the points
\[ \left\{ \left( k, \log \left( |\hat{g}^{ijl}_k| \right) \right) : 0 \leq k \leq M_{jl} (N_i - 1), \ |\hat{g}^{ijl}_k| > 10^{-16} \right\}. \]

In practice, \( \rho_{ijl} \) is roughly the same for all \( 1 \leq i \leq m \) and \( 1 \leq j, l \leq n \) due to the choice of the grid, and we therefore write \( \rho = \rho_{ijl} \). If \( \nu_i < \rho \) and \( N_i \) is sufficiently large, then
\[ \| \hat{g}^{ijl} \|_{\nu_i} \approx \left| \hat{g}^{ijl}_0 \right| \left( 1 + 2 \sum_{k=1}^{M_{jl}(N_i-1)} \left( \nu_i \rho \right)^k \right) \approx \left| \hat{g}^{ijl}_0 \right| \left( \frac{2\rho}{\rho - \nu_i} - 1 \right). \] (3.71)
Altogether, (3.70) and (3.71) yield the constraint

$$
u_i^3 + (\alpha_{ijl} + \rho) \nu_i^2 + (1 - \alpha_{ijl} \rho) \nu_i + \rho \leq 0, \quad \alpha_{ijl} := \frac{2N_i \varepsilon}{L(t_i - t_{i-1})} |\hat{g}_{ijl}^0|.$$ (3.72)

Finally, we determine an interval $[\nu_{\text{min}}, \nu_{\text{max}}] \subset \mathbb{R}_{>0}$ on which (3.72) is satisfied for all $1 \leq i \leq m$ and $1 \leq j, l \leq n$. Then, if $\nu_{\text{max}} > 1$, we choose a weight $\hat{\nu} \in [\nu_{\text{min}}, \nu_{\text{max}}]$ such that $1 < \hat{\nu} < \rho$ and set $\nu_i = \hat{\nu}$ on each subdomain (in practice we set $\hat{\nu} = \frac{1}{2} (\max\{1, \nu_{\text{min}}\} + \nu_{\text{max}})$). If $\nu_{\text{max}} \leq 1$, we increase the number of subdomains (to increase $\rho$) or use a higher number of Chebyshev coefficients $N_i$ and then try again.

### 3.6.1 Lotka-Volterra

We have proven the existence of connecting orbits from $(b, 0, 1 - b, 0)$ to $(1, 0, 0, 0)$ in (3.3) for $a = 5$, $b = \frac{1}{2}$, $D = 3$ and different values of $\kappa$. Recall that these orbits correspond to traveling fronts of (3.2) with wave speed $\kappa$. The choices for these parameter values were somewhat arbitrary and were obtained by experimenting with the parameter values considered in [30]. In particular, we chose the parameters in such a way that the stable eigenvalues associated to $(1, 0, 0, 0)$ consisted of one complex conjugate pair of eigenvalues and one real eigenvalue.

![Figure 3.5: A semi-logarithmic plot of the (nonzero) Chebyshev coefficients of the connecting orbits at $\kappa \in \{-1, -0.7861\}$ on all subdomains for all four components. The black lines correspond to the best line through the points $\left\{ (k, \log |\hat{a}_i^k|_j) \mid |\hat{a}_i^k|_j \geq 10^{-16}, 0 \leq k \leq N_i - 1, 1 \leq i \leq m, 1 \leq j \leq n \right\}$. The results show that the decay rates of the Chebyshev coefficients remained roughly the same for $\kappa \in [-1, -0.7861]$.](image-url)
Connecting orbit at $\kappa = -1$  We started with a numerical approximation of a connecting orbit at $\kappa = -1$ and used the steps outlined in the previous section to obtain the following computational parameters:

- **Parameterization mappings**: we used $K^u = \begin{bmatrix} 13 & 13 \end{bmatrix}$ and $K^s = \begin{bmatrix} 9 & 9 & 9 \end{bmatrix}$
  Taylor coefficients for approximating the local (un)stable manifolds. The length of the stable and unstable eigenvectors was set to $\epsilon_{u,e,k} = 0.0565$ and $\epsilon_{s,e,k} = 0.0635$, respectively. The truncation parameters and the scalings of the eigenvectors were obtained via the procedure in Remark 3.6.1. The scalings of the eigenvectors were relatively small, since the procedure in Remark 3.6.1 was designed to use as little Taylor coefficients as possible to ensure that validation is feasible. If we would allow for larger truncation parameters, the scalings of the eigenvectors (and hence the “size” of the charts on the local (un)stable manifolds) could be increased substantially. However, since it is computationally cheaper to increase the integration time in comparison to increasing the truncation parameters for the (un)stable manifolds, we have chosen to keep the truncation parameters $K^u$ and (especially) $K^s$ small.

- **Chebyshev approximations**: we used $m = 3$ subdomains. The number of Chebyshev modes per subdomain was $N = \begin{bmatrix} 50 & 47 & 50 \end{bmatrix}$. The integration time was set to $L = 15$. The Chebyshev coefficients are shown in Figure 3.5a. This figure shows that the decay rates of the Chebyshev coefficients were approximately the same on each subdomain. Hence the domain decomposition was successful.

- **Validation parameters**: we used $\nu_i = 1.1967$ on each subdomain. This value was obtained from the procedure in Remark 3.6.2. The dimension of the Galerkin projection was $\dim(\mathcal{X}^{NK}) = 5382$. With the above choices for the computational parameters, we successfully validated a connecting orbit at $\kappa = -1$ and proved that the radii-polynomials were negative for $r \in [8.6070 \cdot 10^{-11}, r^*]$. We remark that it is possible to validate the connecting orbit with a smaller number of Chebyshev coefficients as well. The reason for why we used more Chebyshev coefficients than strictly necessary was to get the bounds $Z_{\Pi_i}^{1}$ as small as possible in order to make continuation with large step-sizes feasible.

Non-rigorous continuation  Next, we continued the connecting orbit at $\kappa = -1$ (non-rigorously) by performing pseudo-arc length continuation. At each continuation step, we tried to validate the orbit with the same computational parameters. We succeeded in validating a family of connecting orbits in this way for a finite number of wave speeds $\kappa \in [-1, -0.7861]$, see Figure 3.6a. The reader is referred to the code available at [67] for the exact parameter values $\kappa$ at which the connecting orbits were validated. Here we only give rounded values of $\kappa$ using four decimal places.

We were not able to validate the connecting orbit at the next continuation step $\kappa \approx -0.7767$ with the same computational parameters. The reason for this
Figure 3.6: A bifurcation diagram obtained by continuing the connecting orbit at $\kappa = -1$. The bifurcation curves were computed by performing (non-rigorous) pseudo-arc length continuation. The red points on the curves correspond to validated connecting orbits. In all cases the validation radius $\hat{r}$ was bounded by $2.5347 \cdot 10^{-9}$. The curve in Figure 3.6a was computed with truncation parameters $K^u = [13 \ 13]$, $K^s = [9 \ 9 \ 9]$ and $N = [50 \ 47 \ 50]$. The connecting orbits were validated by using $\nu_i = 1.1967$ on each subdomain. The curve in Figure 3.6b was computed with truncation parameters $K^u = [13 \ 13]$, $K^s = [12 \ 12 \ 12]$ and $N = [55 \ 52 \ 62]$. The connecting orbits were validated by using $\nu_i = 1.1627$ on each subdomain. The “gap” at $\kappa \approx -0.7071$ corresponds to a bifurcation caused by the presence of a resonance at $\kappa = -\frac{1}{2}\sqrt{2}$.

was that the bound for $\Pi^i_Q \left( DF(\hat{x}) - \hat{DF} \right)$ became too large, which was related to the fact that the decay rates of $\hat{q}$ decreased as $\kappa$ increased. In addition, the real part of the stable eigenvalues decreased as well when $\kappa$ increased (see Figure 3.7), which contributed to the deterioration of the bounds for the stable manifold.

On the other hand, the bounds for $\Pi^{ij}_a \left( DF(\hat{x}) - \hat{DF} \right)$ did not deteriorate at all during the continuation process. The main reason for this is that the shape and time parameterization of the orbit remained roughly “the same” throughout the continuation procedure. This caused the decay rates of the Chebyshev coefficients to remain roughly the same as well as shown in Figure 3.5.

To validate connecting orbits for $\kappa \geq -0.7767$, we recomputed the parameterization of the local stable manifold as explained in Remark 3.6.1. This resulted in a parameterization with $K^s = [12 \ 12 \ 12]$ Taylor coefficients. Furthermore, the length of the stable eigenvectors was set to $\epsilon_{s,e_k} = 0.02275$. The resulting parameterization was significantly “smaller” due to the new scaling of the eigenvectors. To ensure that the endpoint of the connecting orbit was contained in the smaller chart, we integrated the connecting orbit forward in time (while keeping the “initial” starting point on the unstable manifold fixed) and increased the integration
time to $L = 20$. We then used the procedures from the previous section again to refine the Chebyshev approximations. In particular, we used $m = 3$ subdomains, $N = [55\ 52\ 62]$ Chebyshev modes and used $\nu_i = 1.1627$ on each subdomain. The dimension of the Galerkin projection was $\dim(\mathcal{X}^N) = 10258$. With these parameter values, we were able to successfully validate a finite number of connecting orbits for $\kappa \in [-0.7767,-0.7075]$, see Figure 3.6b.

Figure 3.6b shows that there is a bifurcation at $\kappa \approx -0.7071$. To understand what caused this bifurcation, we consider the (approximate) stable eigenvalues $\hat{\lambda}_s$ at $\kappa \approx -0.7075$:

$$\hat{\lambda}_s \approx \begin{bmatrix} -0.3537 + 1.541i \\ -0.3537 - 1.541i \\ -0.7072 \end{bmatrix}.$$

Note that $\hat{\lambda}_s^1 + \hat{\lambda}_s^2 \approx \hat{\lambda}_s^3$. This provides numerical evidence for the presence of a resonance at $\kappa \approx -0.7071$ and explains the observed bifurcation. In this relatively simple case, one can prove with pen paper that there is in fact a resonance at $\kappa = -\frac{1}{2}\sqrt{2}$. Hence, in order to validate connecting orbits near $\kappa = -\frac{1}{2}\sqrt{2}$ (and in particular at the resonance point itself), one needs to modify the mapping $F_Q$ by conjugating to a nonlinear normal form (instead of just the linear one) as explained in [75]. Here we do not pursue this issue any further and leave this as a future research project.

To continue the connecting orbit past the resonance, we set $\kappa = -0.7$ and then continued further from this point. We succeeded in validating connecting orbits for $\kappa \in [-0.7,-0.5938]$ without changing the computational parameters, see Figure 3.6b. As $\kappa$ increased, the “size” of the chart on the local stable manifold kept decreasing. As a consequence, for $\kappa > -0.5938$, the endpoint of the connecting orbit was too far away from the stable equilibrium in the sense that the bounds related to the equation $u(1) = Q(\phi)$ were too large (also see Figure 3.8). Although we did
not continue any further, we remark that validation for $\kappa > -0.5938$ is feasible by increasing the integration time and recomputing the Chebyshev approximations as before.

Figure 3.8: The dependence of the stable parameterization variables $\hat{\phi}$ for $\kappa \in [-0.7, -0.5938]$. The results show that $|\text{Re}(\hat{\phi}_1)| = |\text{Re}(\hat{\phi}_2)|$ increased as $\kappa$ increased and eventually became too large (in the sense that the bounds related to the equation $u(1) = Q(\phi)$ were too large). This issue can be resolved by either integrating the connecting orbit forward in time or by enlarging the chart on the local stable manifold in the “directions” of $\phi_1$ and $\phi_2$.

An interesting future research project would be to develop algorithms for automatically detecting when the integration time and/or manifolds need to be modified. We believe that the heuristics in Remarks 3.6.1 and 3.6.2 would be a good starting point for developing such algorithms.

### 3.6.2 Traveling fronts in a fourth order parabolic PDE

We have proven the existence of connecting orbits in (3.5) from $(-1, 0, 0, 0)$ to $(a, 0, 0, 0)$ for $a = -0.1, \kappa = -2$ and various values of $\gamma$. Recall that these orbits correspond to traveling fronts of (3.4) with wave speed $\kappa$. The value for $a$ was obtained from [4] and the wave speed $\kappa$ was chosen through experimentation. We started with a connecting orbit at $\gamma \approx 0.4557$ (see the code available at [67] for the exact numerical data) and then used the procedure as explained at the beginning.
Table 3.1: The (approximate) computational parameters used to validate connecting orbits for $\gamma \in [0.4557, 10.50]$. Each row in the table corresponds to an interval on which we performed (non-rigorous) pseudo-arclength continuation and validated rigorously a finite number of connecting orbits with the same computational parameters. In particular, we used $K^u = [15 \ 15]$ on each interval (though validation with less Taylor coefficients is feasible). We were not able to validate connecting orbits past the right endpoints of the intervals without modifying the computational parameters. In each case, we have indicated the obstruction for validating connecting orbits near the (right) endpoint of the interval.

of this section to select the following computational parameters:

- **Parameterization mappings**: we used $K^u = [15 \ 15]$ and $K^s = [12 \ 12 \ 12]$ Taylor coefficients for approximating the local (un)stable manifolds. The length of the stable and unstable eigenvectors was set to $\epsilon_{u,e,k} = 5.4476 \cdot 10^{-2}$ and $\epsilon_{s,e,k} = 5.3337 \cdot 10^{-3}$, respectively.

- **Chebyshev approximations**: we used $m = 2$ subdomains and $N = [62 \ 61]$ Chebyshev modes. The integration time was set to $L = 30$. The decay rates of the Chebyshev coefficients were approximately the same on each subdomain. Hence the domain decomposition was successful.

- **Validation parameters**: we used $\nu_i = 1.1491$ on each subdomain.

The dimension of the Galerkin projection was $\dim (X_{NK}) = 10314$. With the above choices for the computational parameters, we were able to validate the connecting orbit at $\gamma \approx 0.4557$ and proved that the radii-polynomials were negative for $r \in [4.8332 \cdot 10^{-10}, 2.8332 \cdot 10^{-6}]$.

Next, we performed (non-rigorous) pseudo-arclength continuation and tried to validate the orbits at each continuation step by using the same computational parameters. If the validation failed at a particular continuation step, we determined the cause (as in the previous section) and resolved the issue by modifying the computational parameters. In addition, we also checked in the case of failure whether the dimension of the Galerkin-projection could be significantly reduced by decreasing the integration time and the number of Chebyshev or Taylor coefficients (with special emphasis on reducing the number of Taylor coefficients associated to the stable manifold). The results are summarized in Table 3.1. The corresponding bifurcation curves are shown in Figure 3.9 and the associated traveling wave profiles are depicted in Figure 3.10.
Figure 3.9: Bifurcation diagrams obtained by continuing the connecting orbit at \( \gamma \approx 0.4557 \). The bifurcation curves were computed by performing (non-rigorous) pseudo-arc length continuation. The red points on the curves correspond to validated connecting orbits. In each case, the validation radii were bounded by: (a) \( \hat{r} \leq 4.6324 \cdot 10^{-6} \), (b) \( \hat{r} \leq 5.3757 \cdot 10^{-6} \), (c) \( \hat{r} \leq 9.9475 \cdot 10^{-6} \). The values of the computational parameters are reported in Table 3.1.

We remark that validation of connecting orbits for \( \gamma > 10.50 \) is feasible; the reason for the “obstruction” \( \hat{r} > r^* \) was that the decay rates of the Chebyshev coefficients decreased as \( \gamma \) increased, which eventually resulted in a too large bound for the residual (i.e. the bound \( Y_{a}^{ij} \)). This issue can be easily resolved by using domain decomposition or increasing the number of Chebyshev coefficients (or by just increasing \( r^* \)). Similarly, validation of connecting orbits for \( \gamma < 0.4557 \) (but sufficiently far away from 0) is feasible as well; the bottleneck for small \( \gamma \) is the validation of the local stable manifold. Indeed, as \( \gamma \) decreases, the real parts of the stable eigenvalues decrease (see Figure 3.11) and the number of needed Taylor coefficients increases. An interesting future research project would be to determine how close one can get to the “singular” case \( \gamma = 0 \) with the current method.

Figure 3.10: The first component \( u_1 \), which corresponds to a traveling wave profile of (3.4), of the validated connecting orbits for \( \gamma \in [0.4557, 10.50] \). In each case, for \( \gamma \) close to the left endpoint of the interval, we have colored the associated orbits in dark blue. As \( \gamma \) increased, we used increasingly lighter shades of blue. Note the oscillations for larger values of \( \gamma \).
Figure 3.11: The dependence of the stable and unstable eigenvalues on $\gamma$. 
Validated integration of semilinear parabolic PDEs

4.1 Introduction

During the last decade computer-assisted proofs have become an increasingly effective tool in the study of nonlinear ordinary differential equations and dynamical systems in general. The rapid progress of the development of computer hardware has made it possible to put numerical simulations on a rigorous footing through the construction of theorems whose hypotheses can be verified with the aid of a computer. Today, there exists a large variety of rigorous numerical methods for studying invariant objects, such as equilibria, periodic orbits, connecting orbits, invariant manifolds, etc., in nonlinear ODEs. In particular, we mention the prominent software packages CAPD \cite{2} and COSY \cite{3}. Detailed knowledge of invariant objects can provide deep insight into the global structure of a dynamical system, which for nonlinear systems is typically difficult to obtain from an analysis on paper.

Most of the current methods are applicable to finite dimensional dynamical systems. Only recent attention is devoted to the development of computer-assisted methods for studying invariant objects in dissipative parabolic PDEs, where phase space is infinite dimensional. Such systems arise naturally in physics, e.g. in fluid dynamics, reaction diffusion processes, heat conduction, etc. The term dissipative refers to the fact the long time behavior of these systems is governed by a finite number of degrees of freedom. More precisely, the “interesting” asymptotic behavior often takes place in some finite dimensional manifold, the so-called inertial manifold, on which the dynamics are governed by an ODE, see \cite{85} for instance. Hence, from a mathematical point of view, dissipative parabolic PDEs are a natural class of equations to consider as a first step towards the development of computer-assisted proofs for infinite dimensional dynamical systems.

The development of rigorous numerical methods for the computation of invariant objects in parabolic PDEs is an active field and various methods have already been developed. For example, equilibria of parabolic PDEs have been studied extensively with the aid of a computer in \cite{6,23,27,40,54,61,62,71,86}. Traveling waves in reaction diffusion equations have been studied as well, see \cite{8,16} for instance. In \cite{33,39} computer-assisted methods are developed to validate branches...
of periodic orbits (periodic in both time and space) in the Kuramoto-Shivashinsky equation. This method exploits the fact that the problem can be reformulated into an equivalent fixed-point problem on a space of rapidly decaying sequences by using a Fourier transformation in time and space. An important feature of this approach is that numerical integration of the PDE, which is a computationally expensive and difficult task, is avoided. There are instances, however, in which validated numerical integration is more convenient.

In [63] a computer-assisted procedure is developed to compute local charts on unstable manifolds of hyperbolic equilibria of parabolic PDEs. The authors used their method to prove the existence of a saddle-to-sink connection in Fisher’s equation by verifying that the local unstable manifold of one equilibrium intersected the basin of attraction of another. This approach was successful because they were able to parameterize a sufficiently large patch of the local unstable manifold. If the equilibria were separated “too far” away from each other in phase space, however, it would have become more difficult to directly prove that the unstable manifold and the basin of attraction intersect. In this case, other means, such as rigorous integration of the (semi) flow, are necessary to traverse the distance between the two equilibria. The purpose of this paper is to provide such a tool by developing a rigorous numerical method for validating solutions of dissipative semilinear parabolic PDE.

In this paper we consider parabolic PDEs of the form

$$\frac{\partial u}{\partial t} = \mathcal{L} u + g(u),$$

where $\mathcal{L}$ is a linear differential operator and $g : \mathbb{R} \to \mathbb{R}$ is an arbitrary polynomial. We impose Neumann-boundary conditions and study the case in which

$$\mathcal{L} = \sum_{j=1}^{R} \beta_j \frac{\partial^{2j}}{\partial x^{2j}}, \quad R \in \mathbb{N}, \beta \in \mathbb{R}^R,$$

so that the orthogonal basis generated by the eigenvectors of $\mathcal{L}$ is the Fourier cosine basis. In order to ensure that $\mathcal{L}$ generates a semi-flow, we assume that $\beta_R < 0$ if $R$ is even and $\beta_R > 0$ if $R$ is odd. We shall develop a rigorous numerical method for validating solutions of

$$\begin{cases}
    \frac{\partial u}{\partial t} = \sum_{j=1}^{R} \beta_j \frac{\partial^{2j} u}{\partial x^{2j}} + g(u), & t \in (0, L), \ x \in (0, \pi), \\
    \frac{\partial u}{\partial x} (t, 0) = \frac{\partial u}{\partial x} (t, \pi) = 0, & t \in [0, L], \\
    u (0, x) = f (x), & x \in [0, \pi],
\end{cases} \quad (4.1)$$

where $L > 0$ is a prescribed integration time and $f : \mathbb{R} \to \mathbb{R}$ is a $2\pi$-periodic even function. Since the main motivation for developing a rigorous integrator is to validate invariant objects such as connecting orbits, which are typically extremely
regular solutions of the PDE, we will assume that \( f \) is analytic. We will comment on the significance of these assumptions below.

Before we proceed with a more detailed description of the proposed method, let us mention the work in [7,22,88,90,92] where rigorous integrators for a class of semilinear dissipative PDE are developed. The starting point of these methods is the same as ours; the problem is first reformulated into an infinite dimensional ODE on a sequence space via a Fourier expansion in space. The methodology, however, is very different from the one proposed in this paper. In [88,90,92] the authors develop a rigorous integrator based on the validated integration of a finite dimensional system of ODEs using Lohner-type algorithms developed in [87] and the notion of self-consisted bounds introduced in [92]. This methodology is further developed in [22] and has been used to rigorously study, among other things, periodic orbits in the Kuramoto-Sivashinsky equation [88], globally attracting solutions in the one dimensional Burgers equation [25], heteroclinic connections in the one-dimensional Ohta-Kawasaki model [24] and more. The work in [7] is closer to the approach presented in this paper and is based on reformulating the problem into a preconditioned fixed-point problem by using the variation of constants formula. The fixed-point problem is solved by establishing the existence of a fixed point in a neighborhood of a numerical approximation, obtained through general polynomial interpolation (i.e., the grid-points are free to choose), via Schauder’s Theorem. The fundamental difference with the method presented in this paper is that we use Chebyshev interpolation to perform the numerics and use a functional analytic framework specifically tailored for Chebyshev approximations. Moreover, we use a contraction argument instead of Schauder’s Theorem and thus (in particular) obtain uniqueness of the solution automatically as part of the construction. The two approaches (self-consistent bounds and fixed point formulation) are to a large extent complementary. One advantage of the functional analytic approach in the current paper (as well as [7]) is that it provides a natural setting for extensions to continuation (and bifurcation) studies as well as boundary value problems.

The main idea of this paper is as follows. As mentioned before, our strategy is to first recast (4.1) into an equivalent infinite dimensional system of ODEs by using a Fourier cosine expansion in space. We exploit the fact that the semi-flow generated by \( \mathcal{L} \) is dissipative by integrating the system of ODEs with the aid of the variation of constants formula. In particular, the requirement that \( \beta_R < 0 \) if \( R \) is even and \( \beta_R > 0 \) if \( R \) is odd indicates (through variation of constants) that the long time behavior of the dynamical system is (roughly speaking) governed by only a finite number of Fourier modes, see Section 4.3. This is why it makes sense to study (4.1) via a Fourier expansion in space, on both paper and the computer, in the first place.

The variation of constants formula is used to set up an equivalent zero-finding problem \( F(a) = 0 \) on a space of time-dependent sequences with geometric decay. The reason for considering sequences with geometric decay is that (bounded) solutions of (4.1) are analytic in space. The theory presented in this paper can be adapted to deal with initial conditions \( f \) of class \( C^k \) as well by setting up the zero-finding problem on a space of algebraically decaying sequences, see [47] for
instance. Let us also stress at this point that the method can be used in the case of Dirichlet boundary conditions as well. The only difference in the Dirichlet setting is that we use Fourier sine series (instead of cosine series) and assume that $f$ and $g$ are odd. The set up of the problem (including all the formulae), as explained in Section 4.3, remains exactly the same. The general case with periodic boundary conditions, where $f$ is allowed to be any analytic $2\pi$-periodic function, can be dealt with in an analogous way by considering the “full” Fourier series.

Next, we define a finite dimensional reduction of $F$ by approximating a finite number of time dependent Fourier coefficients with the aid of Chebyshev interpolation. The motivation for using Chebyshev interpolation is twofold. First, interpolation at the Chebyshev points is near to optimal in the sense that the interpolants converge at a near to optimal rate to the objective function under relatively mild smoothness conditions, see [68]. For instance, if the objective function is of class $C^k$, where $k \in \mathbb{N}_0$, and its $(k+1)$-th derivative is of bounded variation, then the interpolation error is of order $m^{-(k+1)}$, where $m$ is the degree of the interpolant. Secondly, Chebyshev interpolation allows for an efficient numerical implementation. For example, interpolants and products of Chebyshev expansions can be computed with the Fast Fourier Transform.

The finite dimensional reduction is used to compute an approximate zero of $F$ with the aid of a computer. The numerical data is then combined with analysis on paper to set up a Newton-like operator $T$ for $F$ based at the approximate zero. Finally, we use pen and paper estimates to derive a finite number of inequalities, which can be rigorously checked with the aid of a computer, to establish that $T$ is a contraction in a small ball centered at the numerical approximation. These inequalities depend on the radius of the ball and part of the computer-assisted method is to determine the admissible radii. This methodology is often referred to as a parameterized Newton-Kantorovich method or the radii-polynomial approach, see [43,84].

The map $T$ is expected to be a contraction for sufficiently small integration times $L > 0$. To perform long time integration we present two alternatives: time-stepping and domain decomposition. The idea of time-stepping is to start with some initial condition and to rigorously integrate the associated initial value problem on a small time interval. If the proof is successful, we compute a rigorous enclosure for the endpoint of the orbit and try to integrate the enclosure forward in time. This process is repeated as long as necessary (and possible). In order for long time integration to be feasible, the size of the enclosures should be “managed” properly. In particular, they should not grow too fast. This is where the dissipativity of the system and the use of high order Chebyshev interpolation come into play. The dissipativity of the system enables us to “control” the tail of the enclosures, while high order Chebyshev interpolation allows us to manage the size of the enclosures for the lower order Fourier coefficients by having small interpolation errors. An alternative to time-stepping is domain decomposition. The idea of this approach is to set up and simultaneously solve a system of ODEs on sufficiently small time intervals. The main focus in this paper is on time-stepping. Nevertheless, we present the details for performing domain decomposition as well, since this is
a powerful generalization of the single step integration technique and it is an attractive proposition in the context of boundary value problems.

Finally, before we present some applications, let us mention a few possible extensions. The theory presented in this paper can be extended in a straightforward manner to deal with systems of parabolic PDEs with polynomial nonlinearities. Furthermore, the theory is easily generalized to deal with higher dimensional rectangular spatial domains. A slightly more involved extension would be to drop the assumption of spatial periodicity by using Chebyshev series instead of Fourier series in the set up of the zero finding problem. The main difficulty to overcome in this case is to incorporate the Neumann-Boundary conditions into the zero finding problem. Another nontrivial extension would be to also allow for more complicated nonlinearities, e.g., nonlinearities of the form $uu_x$. To accomplish this, we would need to incorporate the effect of the exponentials from the variation of constants formulation directly into the tail-estimates for the convolution terms developed in Section 4.4.

All the computations presented in this paper have been implemented in MATLAB using the INTLAB package [66] for interval arithmetic. The code is available at [67].

**Application 4.1.1 (Fisher’s equation).** To test our method, we have validated a solution of Fisher’s equation,

$$
\begin{align*}
\frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + u (1 - u), \quad t \in (0, \tau), \; x \in (0, \pi), \\
\frac{\partial u}{\partial x} (t, 0) &= \frac{\partial u}{\partial x} (t, \pi) = 0, \quad t \in [0, \tau], \\
u (0, x) &= f(x), \quad x \in [0, \pi],
\end{align*}
$$

for $\tau = 8$ time units and $f(x) \approx 2 \exp \left( -\frac{1}{2} \left( x - \frac{\pi}{2} \right)^2 \right)$ (an even periodic analytic extension of an approximate Gaussian centered at $x = \frac{\pi}{2}$). This particular integration time was based on the observation that the orbit seemed to have “converged” to the trivial equilibrium state $u \equiv 1$, see Figure 4.1. The solution was validated by using 400 time steps of length $L = 0.02$. Furthermore, in each time step we used 15th order Chebyshev interpolants to approximate the first ten Fourier coefficients of the solution. The $C^0$-error between the exact and approximate solution was bounded by $1.3658 \cdot 10^{-3}$. The reader is referred to Section 4.9.1 for more details.

**Application 4.1.2 (Swift-Hohenberg).** To illustrate the effectiveness of time-stepping we have validated a “long” orbit in the Swift-Hohenberg equation,

$$
\begin{align*}
\frac{\partial u}{\partial t} &= -\frac{\partial^4 u}{\partial x^4} - 2 \frac{\partial^2 u}{\partial x^2} + (r - 1) u + u^2 - u^3, \quad (t, x) \in (0, \tau) \times (0, \pi), \\
\frac{\partial u}{\partial x} (t, 0) &= \frac{\partial u}{\partial x} (t, \pi) = 0, \quad t \in [0, \tau], \\
u (0, x) &= f(x), \quad x \in [0, \pi],
\end{align*}
$$

145
Figure 4.1: (a), (b), (c) The first three Fourier coefficients \( (a_k)_{k=0}^2 \) of the validated solution of (4.2). (d) The graph of the validated solution \( u \) of (4.3) for \( (t, x) \in [0, 0.4] \times [0, \pi] \). The solution is only depicted for \( t \leq 0.4 \), since the behavior of the orbit for \( t \in [0.4, 8] \) was relatively simple.

for \( \tau = 10 \) time units, \( r = \frac{4}{5} \) and \( f(x) = \frac{1}{10} \cos(x) \). The solution was validated by performing 2000 time steps of length \( L = 0.005 \). In each time-step we used 20th order Chebyshev interpolants to approximate the first five Fourier coefficients of the solution. The \( C^0 \)-error between the exact and numerical approximation was bounded by \( 1.0348 \cdot 10^{-4} \). The validated orbit is depicted in Figure 4.2. The numerical results indicate that the orbit converges to a nontrivial equilibrium state. The reader is referred to Section 4.9.2 for more details.

The paper is organized as follows. In Section 4.2 we introduce notation and provide the necessary background for performing Chebyshev interpolation. In Section 4.3 we set up a fixed point operator for (4.1) and explain how to establish the existence of a fixed point with the aid of a computer. This involves the construction of computable bounds which are developed in full detail in Sections 4.4
Figure 4.2: (a), (b), (c) The first three Fourier coefficients $(a_k)_{k=0}^2$ of the validated solution of (4.3). (d) The graph of the validated solution $u$ of (4.3).

4.5 and 4.6 In addition, we provide details about the numerical implementation in Section 4.7. We explain how to perform long time integration in Section 4.8. Finally, we examine the performance of the proposed method in Section 4.9 by doing some numerical experiments.

4.2 Preliminaries

In this section we introduce notation and provide the necessary background for the tools used in this paper.

Notation for floating point operations Throughout this paper we shall denote the set of floating point numbers on the computer by $\mathbb{F}$. The set of intervals with endpoints in $\mathbb{F}$ is denoted by $\mathbb{IF}$. We denote the up and downward rounding
modes by $\triangle$ and $\triangledown$, respectively, i.e.
\[
\triangle(x) := \min \{ y \in \mathcal{F} : y \geq x \}, \quad \triangledown(x) := \max \{ y \in \mathcal{F} : y \leq x \}, \quad x \in \mathbb{R}.
\]
If $x \in \mathbb{R}^{n_1 \times n_2}$ and $X \in \mathbb{F}^{n_1 \times n_2}$, then we shall write $x \in X$ iff $x_{ij} \in X_{ij}$ for all $1 \leq i \leq n_1$ and $1 \leq j \leq n_2$. Finally, if $x, y \in \mathbb{R}^n$ and $x_k \leq y_k$, $x_k^- = \triangledown(x_k)$, $y_k^+ = \triangle(y_k)$ for $1 \leq k \leq n$, we shall abuse notation and write
\[
\begin{align*}
\triangledown & \left( \begin{bmatrix} x_1, y_1 \\ \vdots \\ x_n, y_n \end{bmatrix} \right) = \begin{bmatrix} x_1^- \\ \vdots \\ x_n^- \end{bmatrix}, \\
\triangle & \left( \begin{bmatrix} x_1, y_1 \\ \vdots \\ x_n, y_n \end{bmatrix} \right) = \begin{bmatrix} y_1^+ \\ \vdots \\ y_n^+ \end{bmatrix}.
\end{align*}
\]

### 4.2.1 Sequence spaces

The functional analytic reformulation of (4.1) in terms of the Fourier cosine coefficients is posed on a space of continuous functions from $[-1, 1]$ into a space of geometrically decaying sequences. To be more precise, we define the space

\[
\ell_1^\nu := \left\{ a \in \mathbb{R}^{N_0} : \sum_{k=0}^{\infty} |a_k| \nu^k < \infty \right\}
\]

endowed with the norm
\[
\|a\|_\nu = |a_0| + 2 \sum_{k=1}^{\infty} |a_k| \nu^k,
\]

where $\nu > 1$ is some prescribed decay-rate to be chosen later.

**Definition 4.2.1** (Space of time-dependent sequences). The space of continuous functions from $[-1, 1]$ into $\ell_1^\nu$ is defined by $\mathcal{X}_\nu := C([-1, 1], \ell_1^\nu)$.

Recall that the Fourier cosine coefficients of the product of two Fourier cosine expansions is given by the symmetric discrete convolution. More precisely, if $u, v : [0, \pi] \to \mathbb{R}$ are Lipschitz, then there exist coefficients $a = (a_k)_{k \in \mathbb{N}_0}, \ b = (b_k)_{k \in \mathbb{N}_0}$ such that
\[
u(x) = a_0 + 2 \sum_{k=1}^{\infty} a_k \cos(kx), \quad v(x) = b_0 + 2 \sum_{k=1}^{\infty} b_k \cos(kx), \quad x \in [0, \pi].
\]

Furthermore,
\[
(u \cdot v)(x) = (a * b)_0 + 2 \sum_{k=1}^{\infty} (a * b)_k \cos(kx), \quad x \in [0, \pi],
\]

where
\[
(a * b)_k := \sum_{k_1 + k_2 = k} a_{|k_1|} b_{|k_2|}, \quad k \in \mathbb{N}_0,
\]

148
see [31] for instance. The motivation for weighing the coefficients \((a_k)_{k \in \mathbb{N}}\) with a factor 2 in (4.4) is that one obtains a convenient bound for \(a \ast b\) in the resulting norm.

**Proposition 4.2.2.** The space \((\ell^n_1, \ast)\) is a commutative Banach algebra.

**Proof.** This follows directly from (4.5) and the triangle inequality. \(\square\)

Another reason for choosing the norm in (4.4) is that bounded linear operators on the resulting space can be analyzed in a rather easy way. To see this, we first introduce the notion of the corner points:

**Definition 4.2.3.** The corner points \(\{\xi_k\}_{k \in \mathbb{N}_0} \subset \ell^n_1\) on the unit ball in \(\ell^n_1\) are defined by \((\xi_k)_l := \varepsilon_k \delta_{kl}\), where \(\delta_{kl}\) is the Kronecker delta function and

\[
\varepsilon_k := \begin{cases} 
1, & k = 0, \\
\frac{2}{\nu} - k, & k \in \mathbb{N}.
\end{cases}
\]

The norm of a bounded linear operator on \(\ell^n_1\) can be computed by simply evaluating it at the corner points as shown in the next proposition:

**Proposition 4.2.4.** Let \((X, \|\cdot\|_X)\) be a normed vector space. If \(L \in \mathcal{B}(\ell_1^n, X)\), then

\[
\|L\|_{\mathcal{B}(\ell_1^n, X)} = \sup_{k \in \mathbb{N}_0} \|L(\xi_k)\|_X.
\]

**Proof.** See Proposition 3.2.8. \(\square\)

Now, suppose \(L \in \mathcal{B}(\ell^n_1, \ell^n_1)\), where \(\nu > 1\). Then \(L\) can be identified with an infinite dimensional matrix, with respect to the canonical Schauder basis \((e_k)_{k \in \mathbb{N}_0}\) for \(\ell^n_1\), where \((e_k)_l = \delta_{kl}\), in the usual way. More precisely, there exists unique coefficients \(\{L_{ij} \in \mathbb{R} : i, j \in \mathbb{N}_0\}\) such that

\[
L(e_j) = \sum_{i=0}^{\infty} L_{ij} e_i \simeq \begin{bmatrix} L_{00} & L_{01} & \cdots \\ L_{10} & L_{11} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \end{bmatrix}, \quad j \in \mathbb{N}_0.
\]

Hence

\[
L(a) = \begin{bmatrix} L_{00} & L_{01} & \cdots \\ L_{10} & L_{11} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \end{bmatrix}, \quad (4.7)
\]

for any \(a \in \ell^n_1\). In this particular setting, Proposition 4.2.4 can be interpreted as the statement that \(\|L\|_{\mathcal{B}(\ell_1^n, \ell_1^n)}\) is a weighted supremum of the \(\ell^n_1\)-norms of the columns of \(L\). Moreover, in this case the converse of Proposition 4.2.4 holds as well:
Proposition 4.2.5. Let $\nu > 1$ and suppose $\{L_{ij} \in \mathbb{R} : i, j \in \mathbb{N}_0\}$ are coefficients such that the expression in (4.7) yields a well-defined linear operator $L : \ell_1^\nu \rightarrow \mathbb{R}^\mathbb{N}_0$, i.e., $(L(a))_k$ is finite for all $a \in \ell_1^\nu$ and $k \in \mathbb{N}_0$. Then $L \in B(\ell_1^\nu, \ell_1^\nu)$ if and only if $\sup_{l \in \mathbb{N}_0} \|L(.l)\|_\nu < \infty$. Moreover, if $L \in B(\ell_1^\nu, \ell_1^\nu)$, then

$$\|L\|_{B(\ell_1^\nu, \ell_1^\nu)} = \sup_{l \in \mathbb{N}_0} \|L(.l)\|_\nu.$$  (4.8)

Proof. See Proposition 3.2.9

4.2.2 Chebyshev interpolation

In this section we recall the basics of Chebyshev interpolation. The reader is referred to [20,64,68] for a comprehensive introduction into the theory of Chebyshev approximation.

Definition 4.2.6 (Chebyshev points). Let $m \in \mathbb{N}$. The $m$-th order Chebyshev points $(t_m^j)_{j=0}^m$ are defined by $t_m^j := \cos\left(\frac{\pi j}{m}\right)$.

Remark 4.2.7. We shall omit the superscript $m$ from the notation whenever it can be easily inferred from the context.

Remark 4.2.8. We have ordered the Chebyshev points from $1$ to $-1$, i.e., $t_0 = 1$ and $t_m = -1$.

Remark 4.2.9. In the literature the points $(t_j^m)_{j=0}^m$ are often referred to as the Chebyshev points of the second kind.

We will refer to the $m$-th order polynomial which interpolates a continuous function $f : [-1, 1] \rightarrow \mathbb{R}$ at the Chebyshev points $(t_j^m)_{j=0}^m$ as the $m$-th order Chebyshev interpolant of $f$. Furthermore, we shall denote the operator which sends a continuous function $f$ to its $m$-th order Chebyshev interpolant by $P_m : C^0[-1, 1] \rightarrow C^0[-1, 1]$. To analyze the errors induced by interpolation at the Chebyshev points we will use the following results:

Theorem 4.2.10. Let $k \in \mathbb{N}_0$. If $f \in C^{k+1}[-1, 1]$, then

$$\|f - P_m(f)\|_{C^0} \leq \Lambda_m \|f^{(k+1)}\|_{C^0},$$

where

$$\Lambda_m := \begin{cases} \frac{\log(m + 1) + \pi}{m + 1}, & k = 0, \\ \frac{8}{\pi k (m - k)^k}, & 1 \leq k \leq m - 1, \\ \frac{1}{2^{m-1}(m + 1)!}, & k = m. \end{cases}$$  (4.9)
Proof. To obtain the estimate for \( k = 0 \), let \( q \) denote the best polynomial approximation of \( f \) of degree \( m \). Then

\[
\| f - P_m(f) \|_{C^0} \leq \| f - q \|_{C^0} + \| q - P_m(f) \|_{C^0}
\]

\[
\leq \left( 1 + \| P_m \|_{B(C^0[-1,1],C^0[-1,1])} \right) \| f - q \|_{C^0},
\]

since \( P_m(q) = q \). Jackson’s Theorem states that the best polynomial approximation of \( f \) of degree \( m \) satisfies the error bound

\[
\| f - q \|_{C^0} \leq \frac{\pi}{2(m + 1)} \| f' \|_{C^0}, \quad f \in C^1[-1,1],
\]

see [20] Chapter 4.5. Furthermore, it is shown in [68] Theorem 15.2 that

\[
\| P_m \|_{B(C^0[-1,1],C^0[-1,1])} \leq \frac{2}{\pi} \log(m + 1) + 1.
\]

Altogether, this proves the result for \( k = 0 \).

The cases \( 1 \leq k \leq m - 1 \) can be found in [68] Theorem 7.2. To obtain the bound for the case \( k = m \), let \( t \in (-1,1) \) be arbitrary and recall the following basic formula for the interpolation error:

\[
f(t) - P_m(f)(t) = f^{(m+1)}(\zeta) \frac{m}{(m + 1)!} \prod_{i=0}^{m} (t - t_i),
\]

for some \( \zeta = \zeta(t) \in (-1,1) \). It follows from the theory developed in [20] Chapter 3 that

\[
\sup_{t \in [-1,1]} \left| \prod_{i=0}^{m} (t - t_i) \right| \leq 2^{1-m}
\]

for the Chebyshev points \( (t_i)_{i=0}^{m} \) (also see [46] for an explicit proof). This proves the result for \( k = m \).

\[\square\]

**Theorem 4.2.11.** Suppose \( f : [-1,1] \to \mathbb{R} \) is analytic. Let \( \mathcal{E}_\rho \subset \mathbb{C} \) denote an open ellipse with foci \( \pm 1 \) to which \( f \) can be analytically extended, where \( \rho > 1 \) is the sum of the semi-major and semi-minor axis of \( \mathcal{E}_\rho \). If \( f \) is bounded on \( \mathcal{E}_\rho \), then

\[
\| f - P_m(f) \|_{C^0} \leq \frac{4\rho^{-m}}{\rho - 1} \sup_{z \in \mathcal{E}_\rho} |f(z)|,
\]

for any \( m \in \mathbb{N} \).

**Proof.** See [68] Theorem 8.2.

\[\square\]

**Remark 4.2.12.** The largest ellipse \( \mathcal{E}_\rho \) to which \( f \) can be analytically extended is often referred to as the Bernstein ellipse associated to \( f \).
Finally, we review methods for computing, differentiating and evaluating Chebyshev interpolants. These operations are used extensively in the numerical implementation and can be performed efficiently in the Chebyshev basis $(T_j)_{j \in \mathbb{N}_0}$:

**Definition 4.2.13.** The Chebyshev polynomials $T_j : [-1, 1] \to \mathbb{R}$ are defined by the relation $T_j (\cos \theta) = \cos (j\theta)$, where $j \in \mathbb{N}_0$ and $\theta \in [0, \pi]$.

**Remark 4.2.14.** In the literature the polynomials $(T_j)_{j \in \mathbb{N}_0}$ are often referred to as the Chebyshev polynomials of the first kind. In particular, the $m$-th order Chebyshev points $(t^m_j)_{j=0}^m$ are the points in $[-1, 1]$ at which $T_m$ attains its extrema. Furthermore, note that $T_j (t^m_k) = \cos \left( \frac{\pi j k}{m} \right)$ for $j \in \mathbb{N}_0$ and $0 \leq k \leq m$.

The polynomials $(T_j)_{j=0}^m$ constitute a basis for the space of $m$-th order polynomials $\mathbb{P}_m$. Hence the $m$-th order Chebyshev interpolant $f_m := P_m(f)$ of a function $f : [-1, 1] \to \mathbb{R}$ can be uniquely written as

$$f_m = \tilde{f}_0^m + 2 \sum_{j=1}^{m} \tilde{f}_j^m T_j. \quad (4.10)$$

We will refer to $(\tilde{f}_j^m)_{j=0}^m$ as the Chebyshev coefficients of order $m$ of $f$ or just simply the Chebyshev coefficients of $f_m$. To avoid clutter in the notation, we will omit the superscript $m$ from the notation whenever it can be easily inferred from the context. We stress, however, that in general $\tilde{f}_{j+1}^m \neq \tilde{f}_j^m$.

**Remark 4.2.15.** Observe that

$$\|f_m\|_{C^0} \leq |\tilde{f}_0| + 2 \sum_{j=1}^{m} |\tilde{f}_j|, \quad (4.11)$$

since $|T_j(t)| \leq 1$ for all $t \in [-1, 1]$ and $j \in \mathbb{N}_0$.

**Remark 4.2.16.** Note that (4.10) is, up to a coordinate transformation, a Fourier cosine series. This is the motivation for using the factor 2 in front of the coefficients $(\tilde{f}_j^m)_{j=1}^m$. In particular, with this convention the Chebyshev coefficients of the product of two Chebyshev expansions is given directly (without a rescaling factor) by the symmetric discrete convolution, see (4.5).

**Computation of a Chebyshev interpolant** Let $m \in \mathbb{N}$. The Chebyshev coefficients of $f_m$ can be computed efficiently with the aid of the Fast Fourier Transform. To explain how to accomplish this, we first recall the definition of the Discrete Fourier Transform:

**Definition 4.2.17 (The Discrete Fourier Transform).** Let $K \in \mathbb{N}$ and set $\omega_K := e^{-2\pi i/K}$. The Discrete Fourier Transform of order $K$ is the map $\text{DFT} : \mathbb{C}^K \to \mathbb{C}^K$.
defined by
\[
[DFT(a)]_j := \sum_{k=0}^{K-1} a_k \omega_k^j, \quad 0 \leq j \leq K - 1. \tag{4.12}
\]

The key observation for computing the Chebyshev coefficients of \(f_m\) is that \((T_j)_j^m\) constitutes an orthogonal basis for \(P_m\) with respect to the inner product
\[
\langle p, q \rangle_m := \sum_{j=1}^{m-1} p(t_j) q(t_j) + \frac{1}{2} (p(t_0) q(t_0) + p(t_m) q(t_m)), \quad p, q \in P_m.
\]

More specifically, a straightforward computation shows that
\[
\langle T_k, T_l \rangle_m = \begin{cases} 
0, & k \neq l, \\
\frac{m}{2}, & k = l \in \{1, \ldots, m - 1\}, \\
m, & k = l \in \{0, m\},
\end{cases}
\tag{4.13}
\]
see [20, Chapter 4.5] for instance.

**Remark 4.2.18.** The relation in \((4.13)\) is often referred to as the discrete orthogonality relation of the Chebyshev polynomials with respect to the Chebyshev points of the second kind.

It follows from the orthogonality relation in \((4.13)\) that the Chebyshev coefficients of \(f_m\) are given by
\[
\tilde{f}_j = \frac{1}{m} \begin{cases} 
\langle f_m, T_j \rangle_m, & 0 \leq j \leq m - 1, \\
\frac{1}{2} \langle f_m, T_j \rangle_m, & j = m.
\end{cases}
\tag{4.14}
\]
We use the definition of the Chebyshev polynomials to see that
\[
\langle f_m, T_j \rangle_m = \sum_{k=1}^{m-1} f_m(t_k) \cos \left( \frac{k j \pi}{m} \right) + \frac{1}{2} (f_m(t_0) + (-1)^j f_m(t_m))
\]
\[
= \frac{1}{2} \cdot DFT \left( [f_m(t_0) \ldots f_m(t_m) f_m(t_{m-1}) \ldots f_m(t_1)]^T \right)_j
\]
for \(0 \leq j \leq m\). Therefore, since \(f_m\) agrees with \(f\) on the \(m\)-th order Chebyshev grid, we conclude that
\[
\tilde{f}_j = \frac{1}{2m} \begin{cases} 
y_j, & 0 \leq j \leq m - 1, \\
\frac{1}{2} y_j, & j = m,
\end{cases}
\tag{4.14}
\]
where
\[
y := DFT \left( [f(t_0) \ldots f(t_m) f(t_{m-1}) \ldots f(t_1)]^T \right).
\]
Differentiation of a Chebyshev interpolant Suppose the Chebyshev coefficients of $f_m$ are explicitly known. Then one can compute the Chebyshev coefficients $(\tilde{f}_j^{'})_{j=0}^{m-1}$ of $f_m'$ by applying a linear transformation to $(\tilde{f}_j)_{j=0}^{m}$. To see how this transformation works, first observe that a primitive of $T_j$ is given by

$$\int T_j = \begin{cases} T_1, & j = 0, \\ \frac{1}{4} (T_2 + T_0), & j = 1, \\ \frac{1}{2} \left( \frac{T_{j+1}}{j+1} - \frac{T_{j-1}}{j-1} \right), & j \in \mathbb{N}_{\geq 2}, \end{cases}$$

which follows directly from the definition of the Chebyshev polynomials. Consequently, a primitive of $f_m'$ is given by

$$\int \left( \tilde{f}_0' + 2 \sum_{j=1}^{m-1} \tilde{f}_j T_j \right) = \frac{1}{2} \tilde{f}_1 T_0 + (\tilde{f}_0 - \tilde{f}_2) T_1 + \sum_{k=2}^{m-1} \left( \frac{\tilde{f}_k' - \tilde{f}_{k+1}'}{k} \right) T_k + \frac{\tilde{f}_{m-1}'}{m} T_m. \quad (4.15)$$

Now, the terms of order one and higher in (4.10) and (4.15) are the same. This observation yields the following recursive relation between the Chebyshev coefficients of $f_m$ and $f_m'$:

$$\begin{align*}
\left\{ \begin{align*}
\tilde{f}_j' & = 2(j+1)\tilde{f}_{j+1} + \tilde{f}_{j+2}', & j = m-1, \ldots, 1, \\
\tilde{f}_j' & = 2\tilde{f}_1 + \tilde{f}_2', & j = 0,
\end{align*} \right. \quad (4.16)
\end{align*}$$

where we have set $\tilde{f}_{m+1}' = \tilde{f}_m' = 0$. Hence we can use the Chebyshev coefficients of $f_m$ to compute the Chebyshev coefficients of $f_m'$.

Evaluation of a Chebyshev interpolant Finally, we recall how to evaluate a Chebyshev interpolant at an arbitrary point $t \in [-1, 1]$ by using the Clenshaw Algorithm. The Clenshaw Algorithm for Chebyshev interpolants is based on the observation that the Chebyshev polynomials satisfy the following recurrence relation:

$$T_j(t) = 2tT_{j-1}(t) - T_{j-2}(t), \quad j \in \mathbb{N}_{\geq 2}, \quad t \in [-1, 1].$$

In fact, if we set $T_1(t) \equiv t$ and $T_0(t) \equiv 1$, then the Chebyshev polynomials $(T_j)_{j \in \mathbb{N}_{\geq 2}}$ can be defined via this recurrence relation.
Now, fix \( t \in [-1, 1] \) and observe that the recurrence relations for \( (T_j(t))_{j=0}^{m} \) are equivalent to

\[
\begin{bmatrix}
1 & -2t & 1 & & \\
-2t & 1 & -2t & 1 & \\
1 & -2t & 1 & & \\
& & \ddots & \ddots & \ddots \\
& & & 1 & -2t & 1
\end{bmatrix}
\begin{bmatrix}
T_0(t) \\
T_1(t) \\
T_2(t) \\
\vdots \\
T_m(t)
\end{bmatrix}
= \begin{bmatrix}
1 \\
-t \\
0 \\
0 \\
0
\end{bmatrix},
\]

where \( H(t) \in \mathbb{R}^{(m+1)\times(m+1)} \) and \( T_m(t), v(t) \in \mathbb{R}^{m+1} \). In particular, note that \( H(t) \) is invertible for any \( t \in [-1, 1] \). Therefore, there exists a unique \( y(t) \in \mathbb{R}^{m+1} \) such that

\[
H(t)^T y(t) = [\tilde{f}_0 \ 2\tilde{f}_1 \ \ldots \ 2\tilde{f}_m]^T. \tag{4.17}
\]

With \( y(t) \) defined via (4.17) we obtain

\[
f_m(t) = [\tilde{f}_0 \ 2\tilde{f}_1 \ \ldots \ 2\tilde{f}_m] T_m(t)
\]

\[
= y^T(t) H(t) T_m(t)
\]

\[
= y^T(t) v(t)
\]

\[
= y_0(t) - ty_1(t). \tag{4.18}
\]

In conclusion, if we can determine \( y_0(t) \) and \( y_1(t) \), then we can compute \( f_m(t) \) by using (4.18). Now, \( y_0(t) \) and \( y_1(t) \) are easily determined by solving the upper triangular system in (4.17), which yields the following recurrence relations:

\[
\begin{cases}
 y_j(t) = 2\tilde{f}_j + 2ty_{j+1}(t) - y_{j+2}(t), & j = m, \ldots, 1, \\
y_j(t) = \tilde{f}_0 + 2ty_1 - y_2, & j = 0,
\end{cases}
\]

where we have set \( y_{m+2}(t) = y_{m+1}(t) = 0 \).

### 4.3 Functional analytic setup

In this section we construct a fixed point map whose fixed points correspond to solutions of (4.1). First, we recast (4.1) into an ODE on \( \ell_1^\nu \) by using a Fourier cosine transformation in space. We then reformulate the ODE as an equivalent zero finding problem on \( \mathcal{X}_\nu \) by using the variation of constants formula. Next, we perform a finite dimensional reduction by approximating a finite number of Fourier modes with the aid of Chebyshev interpolation. This reduction is used to set up a Newton-like map \( T \) based at an approximate (numerically computed) zero. Finally, we derive a finite number of inequalities to establish that \( T \) is a contraction in a neighborhood of the approximate zero.
4.3.1 An equivalent zero-finding problem

In this section we set up a zero finding problem for (4.1) by using the variation of constants formula. To this end, observe that we have the freedom to “redistribute” the linear term \( u \) amongst the differential operator \( L \) and the nonlinearity \( g \). This redistribution can be used, for example, to strengthen the “damping” effect of the exponentials in the variation of constants formula. To make the discussion more precise, let \( \beta_0 \in \mathbb{R} \) be a free parameter (to be chosen later) and rewrite (4.1) as

\[
\frac{\partial u}{\partial t} = \sum_{j=0}^{R} \beta_j \frac{\partial^2 u}{\partial x^{2j}} + g_{\beta_0}(u), \quad g_{\beta_0}(u) := g(u) - \beta_0 u.
\]

Next, let \( (a_k(t))_{k \in \mathbb{N}_0} \) and \( p = (p_k)_{k \in \mathbb{N}_0} \) denote the Fourier cosine coefficients of \( u(t, \cdot) \) and \( f \), respectively, i.e.

\[
u(t, x) = a_0(t) + 2 \sum_{k=1}^{\infty} a_k(t) \cos(kx), \quad f(x) = p_0 + 2 \sum_{k=1}^{\infty} p_k \cos(kx),
\]

for all \( t \in [0, L] \) and \( x \in [0, \pi] \). Define the convolution terms \( c : \ell^1_{\nu} \to \ell^1_{\nu} \) by

\[
g_{\beta_0} \left( a_0 + 2 \sum_{k=1}^{\infty} a_k \cos(kx) \right) = c_0(a) + 2 \sum_{k=1}^{\infty} c_k(a) \cos(kx).
\]

More explicitly,

\[
c(a) = \sum_{j=0}^{N_g} g_j a^j - \beta_0 a,
\]

where \( a^j \) denotes the \( j \)-fold convolution of \( a \) with itself and \( (g_j)_{j=0}^{N_g} \) are the coefficients of \( g \) in the monomial basis.

Substitution of the above expansions into (4.1) yields an infinite dimensional system of ODEs on \([0, L]\) for the Fourier cosine coefficients \( (a_k)_{k \in \mathbb{N}_0} \). The idea is to approximate solutions of this system of ODEs with the aid of Chebyshev interpolation. Therefore, we rescale the time domain \([0, L]\) to \([-1, 1]\), on which the theory of Chebyshev approximations is developed (see Section 4.2.2). Altogether, we arrive at the following system of equations:

\[
\begin{cases}
\frac{da_k}{dt}(t) = \frac{L}{2} \left( \sum_{j=0}^{R} (-1)^j \beta_j k^{2j} a_k(t) + c_k(a(t)) \right), \quad t \in [-1, 1], \\
a_k(-1) = p_k,
\end{cases}
\]

where \( k \in \mathbb{N}_0 \). For notational convenience we shall write

\[
\lambda_k := \frac{L}{2} \sum_{j=0}^{R} (-1)^j \beta_j k^{2j}, \quad k \in \mathbb{N}_0.
\]
In particular, observe that there exists a smallest number \( k^+ \in \mathbb{N}_0 \) such that \( \lambda_{k+1} < \lambda_k < 0 \) for all \( k \geq k^+ \), since \((-1)^R \beta_R < 0\) by assumption. Furthermore, note that \( k^+ \) depends on the choice of \( \beta_0 \).

Finally, integration of (4.20) with the aid of variation of constants yields the following map:

**Definition 4.3.1.** The zero finding map \( F : \mathcal{X}_\nu \to \mathcal{X}_\nu \) for (4.1) is defined by

\[
(F(a)(t))_k := e^{\lambda_k(t+1)} p_k + \frac{L}{2} \int_{-1}^{t} e^{\lambda_k(t-s)} c_k(a(s)) \, ds - a_k(t), \quad k \in \mathbb{N}_0.
\]

**Remark 4.3.2.** While the choice (4.21) is very natural in the context of parabolic equations, especially for large \( k \), the introduction of the “computational” parameter \( \beta_0 \) gives an indication of the flexibility of the setup. As explained above, the modification of \( \lambda_k \to \lambda_k + \beta_0 \) can be absorbed directly into the nonlinearities \( g(u) \) and \( c(a) \). Additionally, in view of the finite dimensional reduction in Section 4.3.2, it could be beneficial to replace a finite set \( \{\lambda_k\}_{k=0}^{N-1} \) by other values, say \( \{\tilde{\lambda}_k\}_{k=0}^{N-1} \). For example, the latter may better reflect the growth and decay of solutions of the finite dimensional truncated system. Such changes would lead to a modification of the nonlinearities, namely \( c_k(a) \to c_k(a) + \left( \lambda_k - \tilde{\lambda}_k \right) a_k \) for \( k = 0, \ldots, N-1 \), which cannot be dealt with in the same way as the “uniform shift” \( \beta_0 \). Instead, one would have to adapt some of the estimates. Although such adaptations are straightforward relative to the comprehensive estimates derived in this paper, we did not pursue such further generalizations here.

**Proposition 4.3.3.** If \( F(a) = 0 \), then

\[
\tilde{u}(t, x) := a_0 \left( \frac{2t}{L} - 1 \right) + 2 \sum_{k=1}^{\infty} a_k \left( \frac{2t}{L} - 1 \right) \cos(kx)
\]

solves (4.1).

**Proof.** Let \( a \in \mathcal{X}_\nu \) be arbitrary and observe that

\[
u(t, x) := a_0(t) + 2 \sum_{k=1}^{\infty} a_k(t) \cos(kx)
\]

converges uniformly to an analytic function on \([0, \pi]\) for any fixed \( t \in [-1, 1] \), since \( a(t) \) decays geometrically to zero (see [31]). In particular, we may compute spatial derivatives (of any order) of \( u \) by differentiating (4.22) term by term.

Next, assume that \( F(a) = 0 \), then \( a_k \in C^1[-1, 1] \) for all \( k \in \mathbb{N}_0 \). Furthermore,

\[
a_k'(t) = \lambda_k a_k(t) + \frac{L}{2} c_k(a(t)), \quad t \in [-1, 1], \quad k \in \mathbb{N}_0,
\]

by construction. Now, let \( 1 < \tilde{\nu} < \nu \) and observe that \( (k^s a_k(t))_{k \in \mathbb{N}_0} \in \ell_{\tilde{\nu}}^1 \) for any \( s \geq 1 \) and \( t \in [-1, 1] \), since \( a(t) \in \ell_\nu^1 \). In fact, it is easy to show that
\[
(a \mapsto (k^*a_k)_{k \in \mathbb{N}_0}) \in B(\mathcal{X}_\nu, \mathcal{X}_\nu). \text{ Hence it follows from (4.23) that } (a'_k(t))_{k \in \mathbb{N}_0} \in \mathcal{X}_\nu.
\]

In particular,
\[
|a'_0(t)| + 2 \sum_{k=1}^{n} |a'_k(t)| \cos (kx) \leq \| (a'_k(t))_{k \in \mathbb{N}_0} \|_{\tilde{p}}, \quad (t, x) \in [-1, 1] \times [0, \pi],
\]
for any \( n \in \mathbb{N} \). Therefore, \( a'_0(t) + 2 \sum_{k=1}^{n} a'_k(t) \cos (kx) \) converges absolutely as \( n \to \infty \) for any \( (t, x) \in [-1, 1] \times [0, \pi] \).

Finally, since \( t \mapsto \| (a'_k(t))_{k \in \mathbb{N}_0} \|_{\tilde{p}} \) is integrable, we may use the Dominated Convergence Theorem to conclude that
\[
a_0(t) + 2 \sum_{k=1}^{n} a_k(t) \cos (kx) \to \int_{t_0}^{t} \left( a'_0(s) + 2 \sum_{k=1}^{\infty} a'_k(s) \cos (kx) \right) ds + u(t_0, x)
\]
as \( n \to \infty \), for any \( x \in [0, \pi] \) and \( t_0, t \in [-1, 1] \) such that \( t \geq t_0 \). Therefore, \( u \) is continuously differentiable with respect to \( t \) and we can compute \( \frac{\partial u}{\partial t} \) by differentiating (4.22) term by term.

Altogether, the above arguments justify the formal computations leading to (4.20) and thus prove the statement.

\[\square\]

### 4.3.2 Finite dimensional reduction

In this section we introduce a finite dimensional reduction of \( F \). To accomplish this we will need to truncate the phase space \( \mathcal{X}_\nu \) and discretize time.

**Definition 4.3.4 (Truncation of phase space).** Let \( N \in \mathbb{N} \) be a truncation parameter. The projection \( \Pi_N : \mathcal{X}_\nu \to C([-1, 1], \mathbb{R}^N) \) is defined by
\[
(\Pi_N (a))_k := \begin{cases} 
a_k, & 0 \leq k \leq N - 1, \\
0, & k \geq N.
\end{cases}
\]

Furthermore, we set \( \Pi_\infty := I - \Pi_N \), where \( I \) is the identity on \( \mathcal{X}_\nu \).

**Remark 4.3.5.** Henceforth we shall identify \( \Pi_N (a) \) with the vector of functions
\[
\begin{bmatrix}
a_0 \\
\vdots \\
a_{N-1}
\end{bmatrix} \in C([-1, 1], \mathbb{R}^N).
\]

**Definition 4.3.6 (Time discretization).** Let \( m \in \mathbb{N} \). The Chebyshev projection \( \Pi_{mN} : \mathcal{X}_\nu \to C([-1, 1], \mathbb{R}^N) \) is defined by
\[
\Pi_{mN} (a) := \begin{bmatrix}
P_m(a_0) \\
\vdots \\
P_m(a_{N-1})
\end{bmatrix},
\]

158
where $P_m : C[-1, 1] \to C[-1, 1]$ is the operator which sends a continuous function to its Chebyshev interpolant (see Section 4.2.2). Furthermore, we set $\Pi_{\infty N} := \Pi_N - \Pi_{mN}$.

We will use the Chebyshev basis $(T_j)_{j \in \mathbb{N}_0}$ (see Definition 4.2.13) to identify $\Pi_{mN}(a)$ with a $N \times (m + 1)$ matrix, or equivalently, an element in $\mathbb{R}^{N(m+1)}$. To be more precise, let $(a_{kj})_{j=0}^m$ denote the Chebyshev coefficients of $[\Pi_{mN}(a)]_k$, i.e.,

$$[\Pi_{mN}(a)]_k = a_{k0} + 2 \sum_{j=1}^{m} a_{kj}T_j, \quad 0 \leq k \leq N - 1,$$

as explained in Section 4.2.2. Then we shall identify $\Pi_{mN}(a)$ with the matrix (or vector)

$$\begin{bmatrix} a_{00} & \cdots & a_{0m} \\ \vdots \\ a_{(N-1)0} & \cdots & a_{(N-1)m} \end{bmatrix} \in \mathbb{R}^{N \times (m+1)} \simeq \mathbb{R}^{N(m+1)}.$$  \hfill (4.24)

Recall that the Chebyshev coefficients of an interpolant can be computed by evaluating the objective function at the Chebyshev grid and applying the Discrete Fourier Transform, see (4.14). For the sake of completeness and introducing notation, we repeat this process again and explain how (4.24) is computed in practice. Let $\varphi_{mN} : X_{\nu} \to \mathbb{R}^{N \times (m+1)}$ denote the map which evaluates the first $N$ Fourier coefficients at the Chebyshev grid, i.e.,

$$\varphi_{mN}(a) := \begin{bmatrix} a_0(t_0) & \cdots & a_0(t_m) \\ \vdots \\ a_{N-1}(t_0) & \cdots & a_{N-1}(t_m) \end{bmatrix}.$$ \hfill (4.25)

Define the map $\mathbf{CT} : \mathbb{R}^{N \times (m+1)} \to \mathbb{R}^{N \times (m+1)}$ by

$$\mathbf{CT}(x) := \frac{1}{2m} \begin{bmatrix} y_{00} & \cdots & y_0(m-1) & \frac{1}{2}y_{0m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{(N-1)0} & \cdots & y_{(N-1)(m-1)} & \frac{1}{2}y_{(N-1)m} \end{bmatrix},$$ \hfill (4.26)

where

$$y_k := \text{DFT} \left( [x_{k0} \cdots x_{km} x_{k(m-1)} \cdots x_{k1}]^T \right), \quad 0 \leq k \leq N - 1,$$

and $\text{DFT}$ is defined in (4.12). Then $\Pi_{mN} = \mathbf{CT} \circ \varphi_{mN}$ by (4.14). The reader is referred to Section 4.2.2 for a more detailed derivation.

The truncation of phase space and discretization of time give rise to the following decomposition of $X_{\nu}$:

$$X_{\nu} = X_{\nu}^{mN} \oplus X_{\nu}^{\infty N} \oplus X_{\nu}^{\infty},$$
where
\[ \mathcal{X}^{mN}_\nu := \Pi_{mN} (\mathcal{X}_\nu), \quad \mathcal{X}^\infty_{\nu N} := \Pi_{\infty N} (\mathcal{X}_\nu), \quad \mathcal{X}^\infty_\nu := \Pi_{\infty} (\mathcal{X}_\nu). \]

We equip \( \mathcal{X}^{mN}_\nu \) with the norm
\[ \|a\|_{\mathcal{X}^{mN}_\nu} := \| \begin{bmatrix} \left| a_{k0} \right| + 2 \sum_{j=1}^{m} \left| a_{kj} \right| \end{bmatrix}^N_{k=0} \|_\nu \]  
(4.27)

by using the identification in (4.24). There are several reasons for choosing this particular norm over the more obvious supremum norm. First of all, the norm in (4.27) is numerically easy to compute, whereas the computation of a supremum norm is relatively complicated. Furthermore, with this norm it is easy to compute operator norms which amounts to computing weighted \( l^1 \)-norms of finite dimensional matrices. This is to be contrasted with the use of a supremum norm, where the analysis of linear operators is much more complicated. Lastly, it follows from (4.11) that the norm in (4.27) is stronger than the supremum norm, i.e.,
\[ \sup_{t \in [-1,1]} \|a(t)\|_\nu \leq \|a\|_{\mathcal{X}^{mN}_\nu}, \quad a \in \mathcal{X}^{mN}_\nu, \]
thereby allowing one to relate the two norms in a straightforward manner.

The subspaces \( \mathcal{X}^\infty_{\nu N} \) and \( \mathcal{X}^\infty_\nu \) are both endowed with the supremum norm, i.e.,
with
\[ \sup_{t \in [-1,1]} \|a(t)\|_\nu. \]

Furthermore, the full space \( \mathcal{X}_\nu \) is equipped with the norm
\[ \|a\|_{\mathcal{X}_\nu} := \max \left\{ \|a\|_{\mathcal{X}^{mN}_\nu}, \epsilon_1^{-1} \|a\|_{\mathcal{X}^\infty_{\nu N}}, \epsilon_2^{-1} \|a\|_{\mathcal{X}^\infty_\nu} \right\}, \]  
(4.28)

where \( \epsilon_1, \epsilon_2 > 0 \) are weights to be chosen later. The purpose of these weights is to provide some control over the truncation errors in phase space and the interpolation errors in time.

Finally, we define a finite dimensional reduction of \( F \):

**Definition 4.3.7 (Finite dimensional reduction of \( F \)).** The finite dimensional reduction \( F_{mN} : \mathcal{X}^{mN}_\nu \rightarrow \mathcal{X}^{mN}_\nu \) of \( F \) is defined by
\[ F_{mN} := \Pi_{mN} \circ F \circ \Pi_{mN} |_{\mathcal{X}^{mN}_\nu}. \]

### 4.3.3 A posteriori analysis

In this section we construct a Newton-like map for \( F \) by using the finite dimensional reduction \( F_{mN} \). To this end, suppose we have computed the following:
(i) An approximate zero \( \hat{a} \in \mathbb{R}^{N(m+1)} \) of \( F_{mN} \).

(ii) The derivative \( DF_{mN}(\hat{a}) \).

(iii) An approximate injective inverse \( A_{mN} \) of \( DF_{mN}(\hat{a}) \).

**Remark 4.3.8.** One can check that \( A_{mN} \) is injective by verifying that the bound in Lemma 4.6.1 is strictly smaller than 1. In fact, our computer-assisted method can only be successful when this inequality is satisfied. Therefore, if the computer-assisted proof is successful, we may a posteriori conclude that \( A_{mN} \) is injective without any further ado.

We will use the finite dimensional data to construct an approximate inverse of \( DF(\hat{a}) \). To this end, let \( h \in \mathcal{X}_\nu \) be arbitrary and observe that

\[
(DF(\hat{a}) h)_k(t) = \frac{L}{2} \int_{-1}^{t} e^{\lambda_k(t-s)} Dc_k(\hat{a}(s)) h(s) \, ds - h_k(t), \quad k \in \mathbb{N}_0, \tag{4.29}
\]

for \( t \in [-1, 1] \). Now, if the order of interpolation is sufficiently large, we expect the interpolation error for \( h_k \) to be relatively large in comparison to the interpolation error for the integral in (4.29), since the integral is of class \( C^1 \) and \( h_k \) is only continuous. Furthermore, if \( N \) is sufficiently large, we expect the integral in (4.29) to be relatively small in comparison to \( h_k \) for \( k \geq N \), since \( \lambda_k = O(k^{2\nu}) \) and \( \lambda_k \to -\infty \) as \( k \to \infty \), see (4.21). For these reasons, we anticipate that

\[
\Pi_{\infty N} DF(\hat{a}) \approx -\Pi_{\infty N}, \quad \Pi_{\infty} DF(\hat{a}) \approx -\Pi_{\infty},
\]

in a small neighborhood of \( \hat{a} \) provided \( m \) and \( N \) are sufficiently large. These observations motivate the following definitions:

**Definition 4.3.9** (Approximation of \( DF(\hat{a}) \)). The approximate derivative \( \hat{DF} : \mathcal{X}_\nu \to \mathcal{X}_\nu \) of \( F \) at \( \hat{a} \) is defined by

\[
\hat{DF} := DF_{mN}(\hat{a}) \oplus (-\Pi_{\infty N}) \oplus (-\Pi_{\infty}).
\]

**Definition 4.3.10** (Approximate inverse of \( DF(\hat{a}) \)). The approximate inverse \( A : \mathcal{X}_\nu \to \mathcal{X}_\nu \) of \( DF(\hat{a}) \) is defined by

\[
A := A_{mN} \oplus (-\Pi_{\infty N}) \oplus (-\Pi_{\infty}).
\]

Next, we define a Newton-like operator \( T \) for \( F \) based at \( \hat{a} \):

**Definition 4.3.11** (Newton-like operator for \( F \)). The Newton-like operator \( T : \mathcal{X}_\nu \to \mathcal{X}_\nu \) for \( F \) based at \( \hat{a} \) is defined by

\[
T := I - AF.
\]

**Remark 4.3.12.** Note that \( F(a) = 0 \) if and only if \( T(a) = a \), since \( A \) is injective.
The idea is to seek fixed points of $T$ in a small neighborhood of $\hat{a}$. To be more precise, let $B_{r,\epsilon}(0)$ denote the closed ball of radius $r > 0$ centered at $0$ in $X_\nu$, i.e.,

$$B_{r,\epsilon}(0) = \left\{ h \in X_\nu : \|\Pi^m_N(h)\|_{X^m_N} \leq r, \quad \|\Pi^\infty_N(h)\|_{X^\infty_N} \leq \epsilon_1 r, \right. \quad \|\Pi^\infty_N(h)\|_{X^\infty} \leq \epsilon_2 r \}.$$  

We shall prove the existence of a fixed point of $T$ in $B_{r,\epsilon}(\hat{a}) = \hat{a} + B_{r,\epsilon}(0)$, where $r > 0$ is an unknown radius to be determined, by using the theorem stated below. A proof of this theorem can be found in [84].

**Theorem 4.3.13** (Contraction mapping principle with variable radius). *Assume the following conditions are satisfied:*

(i) There exist bounds $Y^m_N, Y^\infty_N, Y_\infty \geq 0$ such that

$$\|\Pi^m_N(T(\hat{a}) - \hat{a})\|_{X^m_N} \leq Y^m_N,$$

$$\|\Pi^\infty_N(T(\hat{a}) - \hat{a})\|_{X^\infty_N} \leq Y^\infty_N,$$

$$\|\Pi^\infty(T(\hat{a}) - \hat{a})\|_{X^\infty} \leq Y_\infty,$$

and bounds $Z^m_N(r), Z^\infty_N(r), Z_\infty(r) \geq 0$ such that

$$\sup_{v,w \in B_{1,\epsilon}(0)} \|\Pi^m_N DT(\hat{a} + rv) rw\|_{X^m_N} \leq Z^m_N(r),$$

$$\sup_{v,w \in B_{1,\epsilon}(0)} \|\Pi^\infty_N DT(\hat{a} + rv) rw\|_{X^\infty_N} \leq Z^\infty_N(r),$$

$$\sup_{v,w \in B_{1,\epsilon}(0)} \|\Pi^\infty DT(\hat{a} + rv) rw\|_{X^\infty} \leq Z_\infty(r).$$

(ii) There exists a radius $\hat{r} > 0$ such that

$$Z^m_N(\hat{r}) - \hat{r} + Y^m_N < 0,$$

$$Z^\infty_N(\hat{r}) - \epsilon_1 \hat{r} + Y^\infty_N < 0,$$

$$Z_\infty(\hat{r}) - \epsilon_2 \hat{r} + Y_\infty < 0.$$

*Then $T : B_{\hat{r},\epsilon}(\hat{a}) \rightarrow B_{\hat{r},\epsilon}(\hat{a})$ is a contraction.*

### 4.4 General estimates

In this section we derive estimates which will be used extensively in the computation of the $Y$ and $Z$ bounds as stated in Theorem 4.3.13. We start with deriving a bound for the errors induced by Chebyshev interpolation:
Lemma 4.4.1 (Interpolation-error). Let $j, k \in \mathbb{N}_0$. If $\varphi \in C^j[-1, 1]$, then

$$\left\| (I - P_m) \left( t \mapsto \int_{-1}^{t} e^{\lambda_k(t-s)} \varphi(s) \, ds \right) \right\|_{C^0} \leq \Lambda_{mj} \left[ \lambda_j^k \left| e^{2\lambda_k} - 1 \right| \| \varphi \|_{C^0} + \sum_{i=0}^{j} \lambda_i^k \left\| \varphi^{(j-i)} \right\|_{C^0} \right]$$

for any $j \leq m$, where $\Lambda_{mj}$ is defined in (4.9).

Proof. Let $k \in \mathbb{N}_0$ be arbitrary and observe that the map

$$t \mapsto \int_{-1}^{t} e^{\lambda_k(t-s)} \varphi(s) \, ds$$

is $j + 1$ times continuously differentiable, since $\varphi \in C^j[-1, 1]$. In particular,

$$\left| \frac{d^{j+1}}{dt^{j+1}} \int_{-1}^{t} e^{\lambda_k(t-s)} \varphi(s) \, ds \right| = \begin{cases} \varphi^{(j)}(t), & \lambda_k = 0, \\ \lambda_k^{j+1} \int_{-1}^{t} e^{\lambda_k(t-s)} \varphi(s) \, ds + \sum_{i=0}^{j} \lambda_i^k \varphi^{(j-i)}(t), & \lambda_k \neq 0, \end{cases}$$

for any $t \in [-1, 1]$. Therefore, if $\lambda_k \neq 0$, then

$$\left| \frac{d^{j+1}}{dt^{j+1}} \int_{-1}^{t} e^{\lambda_k(t-s)} \varphi(s) \, ds \right| \leq \lambda_k^{j+1} \| \varphi \|_{C^0} \int_{-1}^{t} e^{\lambda_k(t-s)} \, ds + \sum_{i=0}^{j} \lambda_i^k \left\| \varphi^{(j-i)} \right\|_{C^0}.$$

Moreover, since

$$\int_{-1}^{t} e^{\lambda_k(t-s)} \, ds = \frac{e^{\lambda_k(t+1)} - 1}{\lambda_k} \leq \frac{e^{2\lambda_k} - 1}{\lambda_k}, \quad t \in [-1, 1],$$

we conclude that

$$\left\| t \mapsto \frac{d^{j+1}}{dt^{j+1}} \int_{-1}^{t} e^{\lambda_k(t-s)} \varphi(s) \, ds \right\|_{C^0} \leq \text{sign} (\lambda_k) \left| \lambda_k^j \right| \left( e^{2\lambda_k} - 1 \right) \| \varphi \|_{C^0} + \sum_{i=0}^{j} \lambda_i^k \left\| \varphi^{(j-i)} \right\|_{C^0}.$$

Finally, one easily verifies that the latter estimate holds as well if $\lambda_k = 0$ (and is in fact sharp in this case). The result now follows from Theorem 4.2.10.
Next, we compute a bound for the errors induced by the truncation of phase space. We shall henceforth assume that \( N \geq k^+ \).

**Lemma 4.4.2** (Truncation-error). Define the linear operator \( K_\infty : X_\nu^\infty \to X_\nu^\infty \) by

\[
[K_\infty (a) (t)]_k := \begin{cases} 
0, & 0 \leq k \leq N - 1, \\
\frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} a_k (s) \, ds, & k \geq N,
\end{cases}
\]

then \( K_\infty \) is bounded and

\[
\|K_\infty\|_{B(X_\nu^\infty, X_\nu^\infty)} \leq \frac{L (e^{2\lambda N} - 1)}{2\lambda N}.
\]

**Proof.** Let \( a \in X_\nu^\infty \) and \( t \in [-1, 1] \) be arbitrary, then

\[
\|K_\infty (a) (t)\|_\nu = 2 \sum_{k=N}^{\infty} \frac{L}{2} \left| \int_{-1}^{t} e^{\lambda_k (t-s)} a_k (s) \, ds \right| \nu^k 
\leq \frac{L}{2} \int_{-1}^{t} e^{\lambda N (t-s)} \|a(s)\|_\nu \, ds
\leq \frac{L}{2} \|a\|_{X_\nu^\infty} \int_{-1}^{t} e^{\lambda N (t-s)} \, ds
\leq \frac{L (e^{2\lambda N} - 1)}{2\lambda N} \|a\|_{X_\nu^\infty},
\]

for any \( t \in [-1, 1] \), where in the second line we used the assumption that \( \lambda_k \) is decreasing for \( k \geq N \).

**Remark 4.4.3.** Note that \( \|K_\infty\|_{B(X_\nu^\infty, X_\nu^\infty)} = \mathcal{O} (N^{-2R}) \) as \( N \to \infty \) by (4.21).

Finally, we analyze the map \( Dc (\hat{a}) : \ell^1_\nu \to \ell^1_\nu \), where \( \hat{a} \in \ell^1_\nu \) is an approximate zero of \( F \) obtained through numerical simulation. Recall that \( \hat{a}_k = 0 \) for \( k \geq N \).

The results in this section will be used in Section 4.6 to analyze the difference \( DF (\hat{a}) - \hat{D}F \) between the exact and approximate derivative, which in turn will be used to compute the \( Z \)-bounds as stated in Theorem 4.3.13.

We start with the observation that

\[
Dc (\hat{a}) \, h = \hat{g}_1 \ast h, \quad h \in \ell^1_\nu,
\]

where \( \hat{g}_1 \) are the Fourier cosine coefficients of

\[
x \mapsto g' \left( \hat{a}_0 + 2 \sum_{k=1}^{N-1} \hat{a}_k \cos (kx) \right).
\]

164
In particular, note that \((\tilde{g}_1)_k = 0\) for \(k \geq (N_g - 1)(N - 1) + 1\), where \(N_g\) is the order of \(g\). Therefore, motivated by (4.31), we consider a sequence \(a \in \ell^1_\nu\) such that \(a_k = 0\) for all \(k \geq \tilde{M} := M(N - 1) + 1\), where \(M \in \mathbb{N}\), \(N \in \mathbb{N}_{\geq 2}\), and study the map \(h \mapsto a \ast h\).

**Remark 4.4.4.** In practice we set \(M = N_g - 1\).

The objective in this section is to analyze the operators \(B(a), \Gamma(a) : \ell^1_\nu \rightarrow \ell^1_\nu\) defined by

\[
(B(a)h)_k := \begin{cases} (a \ast h)_k, & 0 \leq k \leq N - 1, \vspace{1mm} \\ 0, & k \geq N, \end{cases}
\]

and

\[
(\Gamma(a)h)_k := \begin{cases} 0, & 0 \leq k \leq N - 1, \vspace{1mm} \\ (a \ast h)_k, & k \geq N. \end{cases}
\]

We start by deriving an explicit matrix representation (as described in (4.7)) for the map \(h \mapsto a \ast h\) on \(\ell^1_\nu\). For the sake of convenience, we first extend \(a\) and \(h\) to bi-infinite sequences, by setting \(a_{-k} := a_k\) and \(h_{-k} = h_k\) for \(k \in \mathbb{N}\), and set up a “bi-infinite” matrix representation for \(h \mapsto a \ast h\). This bi-infinite matrix representation will then be used to derive an “one-sided” matrix representation by using appropriate reflections. To be more precise, observe that

\[(a \ast h)_{k \in \mathbb{Z}} = \begin{bmatrix} a_{-\tilde{M} - 1} & \cdots & a_{-1} & a_{1 - \tilde{M}} \\ \cdots & \ddots & \cdots & \cdots \\ a_{\tilde{M} - 1} & \cdots & a_0 & a_{1 - \tilde{M}} \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ h_{-1} \\ h_0 \\ h_1 \end{bmatrix}.

Here we have identified elements in \(\ell^1_\nu\) with bi-infinite column vectors (with respect to the ordering as depicted above). The shaded regions in grey indicate the position of the “zeroth” row and column.

Next, we convert the above bi-infinite matrix representation to an one-sided representation on \(\mathbb{N}_0\) by “reflecting” all elements on the left hand-side of the zeroth
column to the right. This yields

\[
(a \ast h)_{k \in \mathbb{N}_0} = \begin{bmatrix}
a_0 & \ldots & a_{1-\bar{M}} \\
\vdots & \ddots & \vdots & \ddots \\
a_{\bar{M}-1} & \cdots & a_0 & \cdots & a_{1-\bar{M}} \\
a_{\bar{M}-1} & \cdots & a_0 & \cdots & a_{1-\bar{M}} \\
\vdots & \ddots & \vdots & \ddots & \ddots \\
\end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \vdots \end{bmatrix} + \begin{bmatrix} a_1 & \cdots & a_{\bar{M}-1} \\
\vdots & \ddots & \vdots & \ddots \\
a_{\bar{M}-1} & \cdots & a_0 & \cdots & a_{1-\bar{M}} \\
\vdots & \ddots & \vdots & \ddots & \ddots \\
\end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \end{bmatrix}.
\]  

(4.34)

Altogether, the sum of the above two matrices, which we will denote by \(S(a)\), constitutes an infinite dimensional matrix representation of the linear map \(h \mapsto a \ast h\).

Now, let \(B(a) \in \mathbb{R}^{N \times (\bar{M}+N-1)}\) be the finite dimensional matrix which consists of the first \(N\) rows and \(\bar{M} + N - 1 = (M + 1)(N - 1) + 1\) columns of \(S(a)\). Let \(\Gamma_{\infty}(a)\) denote the infinite dimensional matrix which consists of the rows of \(S(a)\) with index \(N\) and higher. Note that we are using the convention that the indexing of the columns and rows start at zero rather than at one. Then the operators in (4.32) and (4.33) can be represented as

\[
B(a)h = B(a)[h_k]_{k=0}^{\bar{M}+N-2}, \quad \Gamma(a)h = \begin{bmatrix} 0_N \\ \Gamma_{\infty}(a)h \end{bmatrix}.
\]

**Lemma 4.4.5** (Operator norm of \(\Gamma(a)\)). Let \(\Gamma_N(a) \in \mathbb{R}^{(3\bar{M}-1)(N-1) \times (2\bar{M}-1)}\) denote the submatrix \(\{S(a)_{ij} : N \leq i \leq 3(\bar{M}-1), \ 0 \leq j \leq 2(\bar{M}-1)\}\) of \(S(a)\). Then

\[
\|\Gamma(a)\|_{\mathcal{B}(\ell_1^N, \ell_1^N)} = \left\| \begin{bmatrix} 0_{N \times (2\bar{M}-1)} \\ \Gamma_N(a) \end{bmatrix} \right\|_{\mathcal{B}(\ell_1^N, \ell_1^N)}.
\]

**Proof.** It follows directly from the expression in (4.34) that

\[
\|\Gamma(a)\|_{\ell_1^N}^N = \sum_{k=1-\bar{M}}^{\bar{M}-1} |a_k| \nu^k, \quad l \geq 2(\bar{M}-1),
\]

166
where \((\xi_l)_{l \in \mathbb{N}_0}\) are the corner points introduced in Definition 4.2.3. Hence

\[
\|\Gamma(a)\|_{B(\ell_1^2, \ell_1^2)} = \max_{0 \leq l \leq 2(\bar{M} - 1)} \|\Gamma(a)\|_{\nu} = \left\| \begin{bmatrix} 0_{N \times (2\bar{M} - 1)} \\ \Gamma_N(a) \end{bmatrix} \right\|_{B(\ell_1^2, \ell_1^2)}
\]

by Proposition 4.2.5 and the definition of \(\Gamma_N(a)\).

\[\]

**Remark 4.4.6.** The latter result shows that the operator norm of \(\Gamma(a)\) is determined by its first \(2\bar{M} - 1\) columns. In particular, the required computation is finite and can thus be performed with the aid of a computer.

### 4.5 Y-bounds

In this section we compute the bounds for the residual. To this end, observe that

\[
T(\hat{a}) - \hat{a} = -AF(\hat{a}) = (-A_{mN}F_{mN}(\hat{a})) \oplus \Pi_{\infty N}F(\hat{a}) \oplus \Pi_{\infty}F(\hat{a}).
\]

In particular, we immediately see that

\[
Y_{mN} := \|A_{mN}F_{mN}(\hat{a})\|_{X_{\nu}^{mN}}
\]

satisfies the conditions of Theorem 4.3.13. Note that this bound involves a finite number of computations and can thus be implemented on a computer.

The computation of the other two bounds is a bit more involved.

**Proposition 4.5.1 (Computation of \(Y_{\infty N}\)).** Let \(1 \leq d \leq m\) and \(\rho > 1\), then the bound

\[
Y_{\infty N} := \frac{4\rho^{-m}}{\rho - 1} \left\| \begin{bmatrix} e^{\lambda_k (\rho + \rho^{-1}) + \lambda_k p_k} \\ \sum_{k=0}^{N-1} \right\|_{\nu} + \frac{L\Lambda_{md}}{2} \left\| \begin{bmatrix} |\lambda_k| e^{2\lambda_k} - 1 \|c_k(\hat{a})\|_{C^0} + \sum_{i=0}^{d} |\lambda_k^i| \|c_k^{(d-i)}(\hat{a})\|_{C^0} \right\|_{\nu},
\]

where

\[
c_k^{(l)}(\hat{a}) = \frac{d^l}{dt^l} \left( t \mapsto c_k(\hat{a}(t)) \right), \quad l \in \mathbb{N},
\]

and \(0^0 = 1\), satisfies the conditions of Theorem 4.3.13.
Proof. First note that
\[ \Pi_{\infty N} (T (\hat{a}) - \hat{a}) = \Pi_{\infty N} F (\hat{a}) \]
\[ = \Pi_{\infty N} \left( t \mapsto \left[ e^{\lambda_k (t+1) p_k + \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s) c_k (\hat{a} (s))} \, ds} \right]_{k=0}^{N-1} \right) . \quad (4.35) \]
The first terms related to the initial condition are easily analyzed, since
\[ \left\| (I - P_m) \left( t \mapsto e^{\lambda_k (t+1) p_k} \right) \right\|_{C^0} \leq \frac{4\rho^{-m}}{\rho - 1} \sup_{z \in \mathcal{E}_\rho} \left| e^{\lambda_k (z+1) p_k} \right| , \]
for any \( 0 \leq k \leq N - 1 \), by Theorem 4.2.11. Furthermore,
\[ \sup_{z \in \mathcal{E}_\rho} \left| e^{\lambda_k (z+1) p_k} \right| = e^{\lambda_k (\rho+\rho^{-1}) + \lambda_k} |p_k| , \]
since the semi-major axis of \( \mathcal{E}_\rho \) has length \( \rho + \rho^{-1} \). Therefore,
\[ \left\| \Pi_{\infty N} \left( t \mapsto \left[ e^{\lambda_k (t+1) p_k} \right]_{k=0}^{N-1} \right) \right\|_{C^0} \leq \frac{4\rho^{-m}}{\rho - 1} \left\| \left[ e^{\lambda_k (\rho+\rho^{-1}) + \lambda_k} p_k \right]_{k=0}^{N-1} \right\|_{\nu} . \]

To bound the second term in (4.35), let \( 0 \leq k \leq N - 1 \) and observe that
\[ t \mapsto \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s) c_k (\hat{a} (s))} \, ds \quad (4.36) \]
is analytic (even entire). Therefore, we can choose any \( 1 \leq d \leq m \) and use Lemma 4.4.1 to bound the interpolation error associated to (4.36). Altogether, this yields a bound for
\[ \left\| \Pi_{\infty N} \left( t \mapsto \left[ \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s) c_k (\hat{a} (s))} \, ds \right]_{k=0}^{N-1} \right) \right\|_{C^0} , \]
and in turn the stated bound \( Y_{\infty N} \).

Remark 4.5.2. Since (4.36) is entire, we could also have used the bound in Theorem 4.2.11 to bound the interpolation-error. However, this would require one to evaluate (4.36) in the complex plane with the aid of a computer.

Remark 4.5.3 (Implementation). In practice, we do not compute the exact value of \( \left\| c_k^{(l)} (\hat{a}) \right\|_{C^0} \), where \( 0 \leq k \leq N - 1 \) and \( 0 \leq l \leq d \), but instead compute an upper bound for it. To explain how this is accomplished, first observe that \( c_k (\hat{a}) \) is a polynomial of order \( N_g m \), since \( g \) and \( \hat{a} \) are polynomials of order \( N_g \) and
Therefore, we can compute an exact polynomial representation of $c_k(\hat{a})$ by computing its Chebyshev interpolant of order $N_g m$. The Chebyshev coefficients of $c_k^{(l)}(\hat{a})$ for $1 \leq l \leq d$ can then be computed by iterating the procedure in (4.16). Finally, we compute an upper bound for $\|c_k^{(l)}(\hat{a})\|_{C^0}$ by using the estimate in (4.11).

Remark 4.5.4. The bound stated in Proposition 4.5.1 improves whenever $m$ is increased. However, observe that an increase in $d$ (for fixed $m$) does not necessarily yield an improvement unless (roughly speaking) the decrease in $\Lambda_{md}$ outweighs the potential increase in $|\lambda^d|$.

Next, we compute a bound for the residual associated to the truncation in phase space.

Proposition 4.5.5. Suppose that $\|\Pi_\infty(p)\|_\nu \leq \delta_p$, then

$$Y_\infty := L \left( \frac{e^{2\lambda N} - 1}{2\lambda N} \right) \|\Pi_\infty c(\hat{a})\|_{C^0} + \delta_p$$

satisfies the conditions of Theorem 4.3.13.

Proof. First note that

$$\Pi_\infty (T(\hat{a}) - \hat{a}) = \Pi_\infty F(\hat{a}).$$

Next, recall the definition of the operator $K_\infty$ in (4.30) and observe that

$$[\Pi_\infty F(\hat{a}) (t)]_k
\begin{cases}
0, & \text{if } 0 \leq k \leq N - 1, \\
e^{\lambda_k (t+1)} p_k + \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} c_k(\hat{a}(s)) \, ds, & \text{if } N \leq k \leq N_g (N - 1), \\
e^{\lambda_k (t+1)} p_k, & \text{if } k \geq N_g (N - 1) + 1,
\end{cases}$$

for any $t \in [-1, 1]$, since $c_k(\hat{a}) = 0$ for all $k \geq N_g (N - 1) + 1$.

Now, a straightforward computation shows that

$$\left\| h \mapsto \left[ e^{\lambda_k (t+1)} h_k \right]_{k=N}^\infty \right\|_B(\ell_1^N, \ell_1^N) = e^{\lambda N (t+1)},$$

for any $t \in [-1, 1]$. The potential increase in $|\lambda^d|$.
for any fixed $t \in [-1, 1]$, since $\lambda_k$ is decreasing for $k \geq k^+$ and $N \geq k^+$ by assumption. Hence

$$\|\Pi_\infty F(\hat{a})\|_{\mathcal{X}_\nu^\infty} \leq \|K_\infty\|_{\mathcal{B}(\mathcal{X}_\nu^\infty, \mathcal{X}_\omega^\infty)} \|\Pi_\infty c(\hat{a})\|_{\mathcal{X}_\nu^\infty} + \sup_{t \in [-1, 1]} e^{\lambda_N(t+1)} \|\Pi_\infty (p)\|_\nu$$

$$\leq \frac{L(e^{2\lambda_N} - 1)}{2\lambda_N} \|\Pi_\infty c(\hat{a})\|_{\mathcal{X}_\nu^\infty} + \delta_p$$

by Lemma 4.4.2 and the assumption that $\|\Pi_\infty (p)\|_\nu \leq \delta_p$. 

**4.6 Z-bounds**

In this section we compute a bound for the contraction rate of $T$. To this end, let $v, w \in B_{1, \epsilon}(0)$, $r > 0$ be arbitrary and note that

$$DT(\hat{a} + rv)rw = (I - A\overrightarrow{DF})rw - A(DF(\hat{a} + rv) - DF(\hat{a}) + DF(\hat{a}) - \overrightarrow{DF})rw.$$  (4.37)

**4.6.1 First order bounds**

In this section we compute bounds for the first order terms with respect to $r$ in (4.37). Let us start with the easiest term which measures the quality of the approximate inverse and derivative:

**Lemma 4.6.1.** The follow estimate holds for all $h \in B_{1, \epsilon}(0)$:

$$\left\|\Pi_mN (I - A\overrightarrow{DF})h\right\|_{\mathcal{X}_{\nu^m}^m} \leq \left\|I_mN - A_{mN}DF_{mN}(\hat{a})\right\|_{\mathcal{B}(\mathcal{X}_{\nu^m}^m, \mathcal{X}_{\omega^m}^m)}.$$

Moreover,

$$\left\|\Pi_\infty N (I - A\overrightarrow{DF})h\right\|_{\mathcal{X}_{\nu^\infty}^\infty} = 0, \quad \left\|\Pi_\infty (I - A\overrightarrow{DF})h\right\|_{\mathcal{X}_{\nu^\infty}^\infty} = 0.$$

**Proof.** It suffices to observe that

$$\Pi_mN (I - A\overrightarrow{DF}) = I_mN - A_{mN}DF_{mN}(\hat{a}),$$

$$\Pi_\infty N (I - A\overrightarrow{DF}) = 0, \quad \Pi_\infty (I - A\overrightarrow{DF}) = 0,$$

where the last line follows from the fact that $A$ and $\overrightarrow{DF}$ are exact inverses of each other on $\mathcal{X}_{\nu^\infty}^\infty$ and $\mathcal{X}_{\nu^\infty}^\infty$. 

170
Next, we consider the term

\[ A \left( DF(\hat{a}) - \hat{DF} \right) h. \]

We will perform the necessary computations by projecting onto the subspaces \( \mathcal{X}_V^{mN}, \mathcal{X}_V^{\infty} \) and \( \mathcal{X}_V^{\infty} \), respectively. Before we proceed, let us remind the reader that all the numerical computations in this paper are performed with interval arithmetic in order to manage the rounding errors which arise from ordinary floating point arithmetic. Effectively, this means that the numerical algorithms presented in this paper are applied to intervals in \( \mathbb{I} \) instead of floating point numbers. In turn, this allows us to extract mathematically rigorous results from numerical computations on the computer. The reader is referred to Section 4.2 to review the notation introduced for floating point and interval arithmetic.

The result in the next lemma is a typical example of how interval arithmetic can be combined with analysis on paper to obtain a mathematically rigorous result.

**Lemma 4.6.2 (Projection onto \( \mathcal{X}_V^{mN} \)).** Suppose there exist intervals \( I_{kj} \in \mathbb{I} \), for \( 0 \leq k \leq N - 1 \) and \( 0 \leq j \leq m \), such that

\[ L \int_{-1}^{t_j} e^{\lambda_k (t_j - s)} Dc_k (\hat{a}(s)) \left( \Pi_{\infty} + \Pi_{\infty} \right) (h)(s) \, ds \in I_{kj} \]  

(4.38)

for all \( h \in B_{1,\epsilon}(0) \). Then

\[ \left\| \Pi_{mN} A \left( DF(\hat{a}) - \hat{DF} \right) h \right\|_{\mathcal{X}_V^{mN}} \leq \triangle \left( \left\| A_{mN} \cdot \text{CT} (I) \right\|_{\mathcal{X}_V^{mN}} \right), \]

where \( \text{CT} (I) \) is interpreted as a vector in \( \mathbb{I}^{N(m+1)} \).

**Proof.** Let \( h \in B_{1,\epsilon}(0) \) be arbitrary, then

\[
\Pi_{mN} \left( DF(\hat{a}) - \hat{DF} \right) h
\]

\[
= \Pi_{mN} \left( DF(\hat{a}) h - DF_{mN}(\hat{a}) \Pi_{mN}(h) \right)
\]

\[
= \Pi_{mN} DF(\hat{a}) \left( \Pi_{\infty} + \Pi_{\infty} \right) (h)
\]

\[
= \Pi_{mN} \left( t \mapsto \left[ \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} Dc_k (\hat{a}(s)) \left( \Pi_{\infty} + \Pi_{\infty} \right) (h)(s) \, ds \right]_{k=0}^{N-1} \right)
\]

by definition of the approximate derivative \( \hat{DF} \). Next, recall that \( \Pi_{mN} = \text{CT} \circ \varphi_{mN} \), where \( \varphi_{mN} \) and \( \text{CT} \) are defined in (4.25) and (4.26), respectively, see Section 4.3.2. Therefore, since

\[
\varphi_{mN} \left( t \mapsto \left[ \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} Dc_k (\hat{a}(s)) \left( \Pi_{\infty} + \Pi_{\infty} \right) (h)(s) \, ds \right]_{k=0}^{N-1} \right) \in I
\]

171
by assumption, it follows that
\[\Pi_{mN} A \left( DF (\hat{a}) - \hat{DF} \right) h = A_{mN} \Pi_{mN} \left( DF (\hat{a}) - \hat{DF} \right) h \in A_{mN} \cdot CT (I),\]
which proves the result. \(\square\)

**Computation of** \(I\)  
In order for the previous lemma to be of practical use, we need to construct computable intervals \(I_{kj} \subset \mathbb{F}\) which satisfy (4.38). Note that it suffices to set \(I_{km} = \{0\}\) for \(0 \leq k \leq N - 1\), since \(t_m = -1\). To construct enclosures for \(0 \leq j \leq m - 1\), first observe that
\[
\left[ \Pi_{\infty N}(h) \right] (s) \in [-\epsilon_1, \epsilon_1] \times \prod_{k=1}^{N-1} \left[ -\frac{\epsilon_1}{2\nu^k}, \frac{\epsilon_1}{2\nu^k} \right] \times \prod_{k=N}^{\infty} \left[ -\frac{\epsilon_2}{2\nu^k}, \frac{\epsilon_2}{2\nu^k} \right], \quad (4.39)
\]
for any \(s \in [-1, 1]\). Furthermore, recall that
\[\Pi_N DC (\hat{a} (s)) h (s) = B (\hat{g}_1 (s)) \cdot \left[ h_k (s) \right]_{k=0}^{N_g(N-1)-1}, \quad (4.40)\]
where \(\hat{g}_1 (s)\) are the Fourier cosine coefficients of
\[x \mapsto g' \left( \hat{a}_0 (s) + 2 \sum_{k=1}^{N-1} \hat{a}_k (s) \cos (kx) \right),\]
and \(B (\hat{g}_1 (s)) \in \mathbb{R}^{N \times N_g(N-1)}\) is the finite dimensional matrix defined in Section 4.4. Now, the idea is to construct an interval enclosure \(\hat{B}_j \subset \mathbb{F}^{N \times N_g(N-1)}\) for \(B (\hat{g}_1([-1, t_j]))\) and to use this in combination with (4.39), (4.40) to construct \(I_{kj}\).

We start by computing an enclosure for \(\hat{g}_1([-1, t_j])\). First, observe that
\[\hat{g}_1 (s) = \sum_{l=0}^{N_g-1} (l + 1) g_{l+1} \hat{a} (s)^l - \beta_0, \quad (4.41)\]
where \(\hat{a} (s)^l\) denotes the \(l\)-fold convolution of \(\hat{a} (s)\) with itself and \((g_l)_{l=1}^{N_g}\) are the coefficients of \(g\) in the monomial basis. In particular, \((\hat{g}_1)_l = 0\) for \(l \geq (N_g - 1)(N - 1) + 1\). The formula in (4.41) shows that \(\hat{g}_1\) is a polynomial of order \((N_g - 1) m\) in \(s\), since \(\hat{a}\) is a polynomial of order \(m\). Therefore, for each \(0 \leq l \leq (N_g - 1)(N - 1)\), we can compute an exact polynomial representation of \((\hat{g}_1)_l\) by computing its Chebyshev interpolant of order \((N_g - 1) m\). This interpolant can then be used to determine an enclosure for \((\hat{g}_1([-1, t_j]))_l\). More precisely, we construct intervals \(G_{lj} \subset \mathbb{F}\) such that
\[(\hat{g}_1([-1, t_j]))_l \subset G_{lj}, \quad 0 \leq l \leq (N_g - 1)(N - 1), \quad (4.42)\]
by evaluating the Chebyshev interpolant of \((\hat{g}_1)_i\) on \([-1, t_j]\) with the Clenshaw Algorithm, see (4.18), using a sufficiently fine subdivision of the interval.

Next, we set

\[
\tilde{B}_j := B\left([G_{0j} \ldots G_{(N_g-1)(N-1)j}]^T\right)
\]

(4.43)

and construct interval enclosures \(e_{kj}, L, t_j \in \mathbb{IF}, h \in \mathbb{IF}^{N_g(N-1)}\) for \(e^{\lambda_k[0,1+t_j]}\), \(L, t_j\) and

\[
[-\epsilon_1, \epsilon_1] \times \prod_{k=1}^{N-1} \left[ \frac{-\epsilon_1}{2\nu_k\nu} \frac{\epsilon_1}{2\nu_k\nu} \right] \times \prod_{k=N}^{N_g(N-1)} \left[ \frac{-\epsilon_2}{2\nu_k\nu} \frac{\epsilon_2}{2\nu_k\nu} \right],
\]

respectively. It then follows from (4.39) and (4.40) that

\[
e^{\lambda_k(t_j-s)}Dc_k(\hat{a}(s)) (\Pi_{\infty,N} + \Pi_{\infty})(h)(s) \in e_{kj} \tilde{B}_j h, \quad s \in [-1, t_j],
\]

for any \(h \in B_{1,\epsilon}(0)\). Hence it suffices to set \(I_{kj} := [\mathcal{I}_{kj}^-, \mathcal{I}_{kj}^+]\), where

\[
\mathcal{I}_{kj}^- := \nabla \left( \frac{L(t_j + 1)}{2} \nabla \left( e_{kj} \tilde{B}_j h \right) \right), \quad \mathcal{I}_{kj}^+ := \Delta \left( \frac{L(t_j + 1)}{2} \Delta \left( e_{kj} \tilde{B}_j h \right) \right).
\]

(4.44)

**Remark 4.6.3.** In practice \(\mathcal{I}_{kj}^- \approx -\mathcal{I}_{kj}^+\), since \(h\) consists of intervals of the form \([-y, y]\).

Next, we consider the bounds related to the errors induced by interpolation. For notational convenience, we shall denote the operator norm of a functional \(f \in (\ell^1_\nu)^*\) by \(\|f\|_{\nu}^*\) and write \(\tilde{c} := 1 + \epsilon_1 + \epsilon_2\).

**Lemma 4.6.4** (Projection onto \(X_{\nu,\infty}^N\)). The following estimate holds for any \(h \in B_{1,\epsilon}(0)\):

\[
\left\| \Pi_{\infty,N} A \left( DF(\hat{a}) - \hat{DF}\right) h \right\|_{X_{\nu,\infty}^N} \leq \frac{L \tilde{c} \log (m+1) + \pi}{m+1} \left\| \left( e^{2\lambda_k} - 1 \right) \max_{s \in [-1,1]} \|Dc_k(\hat{a}(s))\|_{\nu}^* \right\|_{\nu}^{N-1}\right\|_{\nu}.
\]

**Proof.** Let \(h \in B_{1,\epsilon}(0)\) be arbitrary and observe that

\[
\Pi_{\infty,N} \left( DF(\hat{a}) - \hat{DF} \right) h = \Pi_{\infty,N} (DF(\hat{a}) h + h)
\]

\[
= \Pi_{\infty,N} \left( t \mapsto \left[ L \int_{-1}^t e^{\lambda_k(t-s)} Dc_k(\hat{a}(s)) h(s) \ ds \right]_{k=0}^{N-1} \right).
\]

173
Next, let $0 \leq k \leq N - 1$ and use Lemma 4.4.1 to see that

$$
\left\| (I - P_m) \left( t \mapsto \frac{L}{2} \int_{-1}^{t} e^{\lambda_k(t-s)} Dc_k (\hat{a}(s)) h(s) \, ds \right) \right\|_{C^0}
\leq \frac{L \log (m+1) + \pi}{m+1} \left( \left| e^{2\lambda_k} - 1 \right| + 1 \right) \| s \mapsto Dc_k (\hat{a}(s)) h(s) \|_{C^0}.
$$

Moreover,

$$
\| s \mapsto Dc_k (\hat{a}(s)) h(s) \|_{C^0} \leq \tilde{\epsilon} \max_{s \in [-1,1]} \| Dc_k (\hat{a}(s)) \|_{\nu}^*,
$$

since $h \in B_{1,\epsilon}(0)$, which proves the result.

**Remark 4.6.5 (Implementation).** The $k$-th row $\tilde{b}_k \in \mathbb{F}^{1 \times N_{\nu}(N-1)}$ of the (interval-valued) matrix $\tilde{B}_0$ defined in (4.43) is an enclosure for $Dc_k (\hat{a} [-1,1])$, see (4.40). Hence

$$
\max_{s \in [-1,1]} \| Dc_k (\hat{a}(s)) \|_{\nu}^* \leq \triangle \left( \left\| \tilde{b}_k \right\|_{\nu}^* \right), \quad 0 \leq k \leq N - 1.
$$

Furthermore, $\triangle \left( \left\| \tilde{b}_k \right\|_{\nu}^* \right)$ can be computed with the aid of a computer by using interval arithmetic and the result in Proposition 4.2.5. In practice, we use these bounds to implement the estimate stated in Lemma 4.6.4.

Finally, we consider the bound related to the truncation error.

**Lemma 4.6.6 (Projection onto $X_{\nu}^{\infty}$).** The following estimate holds for any $h \in B_{1,\epsilon}(0)$:

$$
\left\| \Pi_{\infty} A \left( DF (\hat{a}) - \hat{DF} \right) h \right\|_{X_{\nu}^{\infty}} \leq \frac{L (e^{2\lambda_N} - 1)}{2\lambda_N} \tilde{\epsilon} \max_{s \in [-1,1]} \| \Gamma (\hat{g}_1(s)) \|_{\mathcal{B}(\ell^1,\ell^1)},
$$

where $\Gamma (\hat{g}_1(s))$ is defined in (4.33).

**Proof.** As before, we start by computing

$$
\Pi_{\infty} \left( DF (\hat{a}) - \hat{DF} \right) h = \Pi_{\infty} \left( DF (\hat{a}) h + h \right)
= \Pi_{\infty} \left( t \mapsto \left[ \frac{L}{2} \int_{-1}^{t} e^{\lambda_k(t-s)} Dc_k (\hat{a}(s)) h(s) \, ds \right]_{k \in \mathbb{N}_0} \right)
= K_{\infty} \Pi_{\infty} \left( s \mapsto Dc (\hat{a}(s)) h(s) \right).
$$
Next, observe that
\[
\| K_\infty \Pi_\infty (s \mapsto Dc (\hat{a} (s)) h(s)) \|_{\mathcal{X}_\nu^\infty}
\leq \| K_\infty \|_{\mathcal{B}(\mathcal{X}_\nu^\infty, \mathcal{X}_\nu^\infty)} \| \Pi_\infty (s \mapsto Dc (\hat{a} (s)) h(s)) \|_{\mathcal{X}_\nu^\infty}
\leq \frac{L (e^{2\lambda N} - 1)}{2\lambda N} \sup_{s \in [-1, 1]} \| \Pi_\infty Dc (\hat{a} (s)) h(s) \|_{\nu}
\leq \frac{L (e^{2\lambda N} - 1)}{2\lambda N} \bar{\epsilon} \sup_{s \in [-1, 1]} \| \Pi_\infty Dc (\hat{a} (s)) \|_{\mathcal{B}(\ell^1_\nu, \ell^1_\nu)} ,
\]
by Lemma 4.4.2 and since \( h \in B_{1, \epsilon} (0) \). Finally, observe that
\[
\| \Pi_\infty Dc (\hat{a} (s)) \|_{\mathcal{B}(\ell^1_\nu, \ell^1_\nu)} = \| \Gamma (\hat{g}_1 (s)) \|_{\mathcal{B}(\ell^1_\nu, \ell^1_\nu)} , \quad s \in [-1, 1],
\]
by (4.31) and the definition of \( \Gamma (\hat{g}_1 (s)) \), which proves the claim. \( \square \)

**Remark 4.6.7 (Implementation).** Observe that
\[
\| \Gamma (\hat{g}_1 (s)) \|_{\mathcal{B}(\ell^1_\nu, \ell^1_\nu)} = \left\| \begin{bmatrix} 0_{N \times (2(N_g - 1)(N - 1) + 1)} & 0_{N \times N} \\ \Gamma_N (\hat{g}_1 (s)) & \end{bmatrix} \right\|_{\mathcal{B}(\ell^1_\nu, \ell^1_\nu)} , \quad s \in [-1, 1], \quad (4.45)
\]
by Lemma 4.4.5. Furthermore, the operator norm in (4.45) can be explicitly computed by using Proposition 4.2.5. In particular, note that this computation is finite (for fixed \( s \in [-1, 1] \)). Now, recall the definition of the enclosures \( \mathcal{G}_{l_i} \) in (4.42) and observe that \((\hat{g}_1([-1,1]))_l \subset \mathcal{G}_{l_0} \in \mathbb{IF} \) for \( 0 \leq l \leq (N_g - 1)(N - 1) \). Therefore, we can construct an enclosure for \( \Gamma_N (\hat{g}_1 [-1, 1]) \) by computing the interval-valued matrix
\[
\Gamma_N \left( \begin{bmatrix} \mathcal{G}_{l_0} & \cdots & \mathcal{G}_{(N_g - 1)(N - 1)0} \end{bmatrix}^T \right).
\]
In practice, we use this enclosure to compute an upper bound for
\[
\max_{s \in [-1, 1]} \| \Gamma (\hat{g}_1 (s)) \|_{\mathcal{B}(\ell^1_\nu, \ell^1_\nu)} .
\]

### 4.6.2 Second order bounds

In this section we compute the second order bounds with respect to \( r \) for (4.37). A straightforward computation shows that
\[
[(DF (\hat{a} + rv) - DF (\hat{a})) h]_k (t) = \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} \left( Dc_k (\hat{a} (s) + rv (s)) - Dc_k (\hat{a} (s)) \right) h(s) \, ds
\]
\[
= \frac{Lr}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} \left( \int_{0}^{1} D^2 c_k (\hat{a} (s) + \tau rv (s)) [h (s), v (s)] \, d\tau \right) ds \quad (4.46)
\]
for any \( k \in \mathbb{N}_0 \), by the (generalized) Mean Value Theorem.

**Lemma 4.6.8.** Let \( r^* > 0 \) be an a priori upper bound for the radius \( r > 0 \), then

\[
\left\| D^2 c(\hat{a}(s) + \tau rv(s)) [h(s), v(s)] \right\|_\nu \leq \tilde{\epsilon}^2 \tilde{g}''(\|\hat{a}(s)\|_\nu + r^* \tilde{\epsilon})
\]

for any \( s \in [-1, 1] \), \( \tau \in [0, 1] \) and \( h, v \in B_{1, \epsilon}(0) \), where \( \tilde{g} : \mathbb{R} \to \mathbb{R} \) is defined by

\[
\tilde{g}(x) := \sum_{j=0}^{N_g} g_j x^j.
\]

**Proof.** Let \( s \in [-1, 1] \), \( \tau \in [0, 1] \), \( h, v \in B_{1, \epsilon}(0) \) be arbitrary, then

\[
D^2 c(\hat{a}(s) + \tau rv(s)) [h(s), v(s)]
\]

\[
= \sum_{j=0}^{N_g-2} (j+2)(j+1)g_{j+2} (\hat{a}(s) + \tau rv(s))^j * h(s) * v(s)
\]

by (4.19). Now use the Banach algebra estimate to see that

\[
\left\| D^2 c(\hat{a}(s) + \tau rv(s)) [h(s), v(s)] \right\|_\nu
\]

\[
\leq \tilde{\epsilon}^2 \sum_{j=0}^{N_g-2} (j+2)(j+1) |g_{j+2}| (\|\hat{a}(s)\|_\nu + r^* \tilde{\epsilon})^j \|h(s)\|_\nu \|v(s)\|_\nu
\]

where in the last line we used that \( r \leq r^* \), \( \tau \leq 1 \) and \( h, v \in B_{1, \epsilon}(0) \).

We will use the preceding result extensively to compute the second order bounds. Before we proceed, let us recall that the weights \((\varepsilon_k)_{k \in \mathbb{N}_0}\) were introduced in (4.6).

**Lemma 4.6.9** (Projection onto \( X_{\nu}^{mN} \)). Let \( 0 \leq j \leq m \) and \( 0 \leq k \leq N-1 \). Choose \( J^+_{kj} \in \mathbb{I}^N_\mathbb{I} \) such that

\[
J^+_{kj} \geq \frac{L}{2} (1 + t_j) \tilde{\epsilon}^2 \varepsilon_k \max_{s \in [-1, t_j]} e^{\lambda_k(t_j-s)} \tilde{g}''(\|\hat{a}(s)\|_\nu + r^* \tilde{\epsilon})
\]

Set \( J_{kj} := [-J^+_{kj}, J^+_{kj}] \) and define \( J \in \mathbb{I}^{N \times (m+1)} \) by

\[
J := \begin{bmatrix}
J_{00} & \ldots & J_{0m} \\
\vdots & \ddots & \vdots \\
J_{(N-1)0} & \ldots & J_{(N-1)m}
\end{bmatrix},
\]

176
then
\[ \| \Pi_{mN} A(DF(\hat{a} + rv) - DF(\hat{a})) h \|_{\mathcal{X}^\nu_{\mu N}} \leq \sup \| A_{mN} \cdot CT(\mathcal{J}) \|_{\mathcal{X}^\nu_{\mu N}} r \]
for any \( h \in B_{1,\epsilon}(0) \).

**Proof.** Let \( h \in B_{1,\epsilon}(0) \) be arbitrary, define \( \eta_k : [-1, 1] \to \mathbb{R} \) by
\[ \eta_k(s) := \int_0^1 D^2 c_k(\hat{a}(s) + \tau rv(s)) [h(s), v(s)] d\tau, \]
for \( 0 \leq k \leq N - 1 \), and observe that
\[ \Pi_{mN} (DF(\hat{a} + rv) - DF(\hat{a})) h = \Pi_{mN} \left( t \mapsto \left[ \frac{Lr}{2} \int_{-1}^t e^{\lambda_k(t-s)} \eta_k(s) \, ds \right]_{k=0}^{N-1} \right) \]
by (13). In particular,
\[ \left| \frac{L}{2} \int_{-1}^{t_j} e^{\lambda_k(t_j-s)} \eta_k(s) \, ds \right| \]
\[ \leq \frac{L}{2} \int_{-1}^{t_j} e^{\lambda_k(t_j-s)} \sup_{\tau \in [0,1]} |D^2 c_k(\hat{a}(s) + \tau rv(s)) [h(s), v(s)]| \, ds \]
\[ \leq \mathcal{J}_{kj}^+ \]
by Lemma 4.6.8 Hence
\[ \varphi_{mN} \left( t \mapsto \left[ \frac{L}{2} \int_{-1}^t e^{\lambda_k(t-s)} \eta_k(s) \, ds \right]_{k=0}^{N-1} \right) \in \mathcal{J}. \]
Consequently,
\[ \Pi_{mN} A(DF(\hat{a} + rv) - DF(\hat{a})) h \]
\[ = A_{mN} \Pi_{mN} (DF(\hat{a} + rv) - DF(\hat{a})) h \in A_{mN} \cdot CT(\mathcal{J}) r, \]
which proves the claim. \( \square \)

**Lemma 4.6.10** (Projection onto \( \mathcal{X}^{\nu\infty N} \)). The following estimate holds for any \( h \in B_{1,\epsilon}(0) \):
\[ \| \Pi_{\infty N} A(DF(\hat{a} + rv) - DF(\hat{a})) h \|_{\mathcal{X}^{\nu\infty N}} \]
\[ \leq \frac{L}{2} \Lambda m_0 \epsilon^2 \max_{s \in [-1,1]} \hat{g}'' \left( \| \hat{a}(s) \|_\nu + \epsilon^* \right) \left( N + \sum_{k=0}^{N-1} |e^{2\lambda_k} - 1| \right) r. \]
Proof. First note that
\[ \Pi_\infty A (DF (\hat{a} + rv) - DF (\hat{a})) h = -\Pi_\infty (DF (\hat{a} + rv) - DF (\hat{a})) h. \]
Furthermore,
\[ \left\| (I - P_m) \left( t \mapsto \frac{L}{2} \int_{-1}^{t} e^{\lambda_k (t-s)} \int_{0}^{1} D^2 c_k (\hat{a} (s) + \tau rv (s)) [h (s), v (s)] d\tau \, ds \right) \right\|_{C^0} \]
\[ \leq \frac{L}{2} \Lambda_{m_0} \left( |e^{2\lambda_k} - 1| + 1 \right) \left\| s \mapsto \int_{0}^{1} D^2 c_k (\hat{a} (s) + \tau rv (s)) [h (s), v (s)] d\tau \right\|_{C^0} \]
\[ \leq \frac{L}{2} \Lambda_{m_0} \left( |e^{2\lambda_k} - 1| + 1 \right) \sup_{s \in [-1,1]} \sup_{\tau \in [0,1]} |D^2 c_k (\hat{a} (s) + \tau rv (s)) [h (s), v (s)]| \]
\[ \leq \frac{L}{2} \Lambda_{m_0} \left( |e^{2\lambda_k} - 1| + 1 \right) \tilde{\epsilon}^2 \epsilon_k \max_{s \in [-1,1]} \tilde{g}'' (\|\hat{a} (s)\|_\nu + r^* \tilde{\epsilon}) , \]
for any \( 0 \leq k \leq N - 1 \) by Lemmas 4.4.1 and 4.6.8. Finally, we take the \( \ell_1^1 \)-norm of the latter quantity (for \( 0 \leq k \leq N - 1 \)) to obtain the result.

Lemma 4.6.11 (Projection onto \( X^\infty_\nu \)). The follow estimate holds for any \( h \in B_{1,\epsilon} (0) \):
\[ \| \Pi_\infty A (DF (\hat{a} + rv) - DF (\hat{a})) h \|_{X^\infty_\nu} \]
\[ \leq \frac{L (e^{2\lambda_N} - 1)}{2\lambda_N} \tilde{\epsilon}^2 \epsilon_k \max_{s \in [-1,1]} \tilde{g}'' (\|\hat{a} (s)\|_\nu + r^* \tilde{\epsilon}) \cdot r. \]

Proof. It follows directly from (13) that
\[ \Pi_\infty A (DF (\hat{a} + rv) - DF (\hat{a})) h \]
\[ = -\Pi_\infty (DF (\hat{a} + rv) - DF (\hat{a})) h \]
\[ = -K_\infty \Pi_\infty \left( s \mapsto \int_{0}^{1} D^2 c (\hat{a} (s) + \tau rv (s)) [h (s), v (s)] d\tau \right) r. \]
Hence
\[ \| \Pi_\infty A (DF (\hat{a} + rv) - DF (\hat{a})) h \|_{X^\infty_\nu} \]
\[ \leq \| K_\infty \|_{B(X^\infty_\nu, X^\infty_\nu)} \left\| s \mapsto \int_{0}^{1} D^2 c (\hat{a} (s) + \tau rv (s)) [h (s), v (s)] \right\|_{X^\infty_\nu} r. \]
The result now follows from Lemmas 4.4.2 and 4.6.8.
Finally, we are ready to construct the \( Z \)-bounds. It follows from Lemmas 4.6.1, 4.6.2, 4.6.9 and the decomposition in (4.37) that
\[
Z_{mN}(r) := \Delta \left( \| I_{mN} - A_{mN} DF_{mN}(\hat{a}) \|_{\mathcal{B}(\mathcal{X}_m^{\infty}, \mathcal{X}_m^{\infty})} + \| A_{mN} \cdot \mathbf{CT} (I) \|_{\mathcal{X}_m^{\infty}} \right) r
\]
\[+ \Delta \left( \| A_{mN} \cdot \mathbf{CT} (J) \|_{\mathcal{X}_m^{\infty}} \right) r^2
\]
satisfies the conditions of Theorem 4.3.13. Similarly, if we denote the bounds in Lemmas 4.6.4, 4.6.6, 4.6.10, 4.6.11 by \( Z^1_{\infty N}, Z^2_{\infty N}, Z^3_{\infty N}, Z^4_{\infty} \), respectively, then
\[
Z_{\infty N}(r) := Z^1_{\infty N} r + Z^2_{\infty N} r^2, \quad Z_{\infty}(r) := Z^3_{\infty} r + Z^4_{\infty} r^2
\]
satisfy the conditions of Theorem 4.3.13.

4.7 Implementation details

In this section we explain how to compute rigorous enclosures for
\[
\int_{-1}^{t_j} e^{\lambda_k(t_j - s)} \psi(s) \, ds
\]
\[= \frac{t_j + 1}{2} \int_{-1}^{1} \exp \left( \frac{\lambda_k}{2} (t_j + 1)(1 - s) \right) \psi \left( \frac{t_j + 1}{2} (s + 1) - 1 \right) \, ds, \quad (4.47)
\]
where \( \psi : [-1, 1] \to \mathbb{R} \) is given by
\[
\psi = \psi_0 + 2 \sum_{l=1}^{M} \psi_l T_l,
\]
for some \( M \in \mathbb{N}_0, 0 \leq k \leq N - 1 \) and \( 0 \leq j \leq m - 1 \). In practice, we set \( \psi = c_k(\hat{a}) \) and \( M = N_g m \). These enclosures are needed to rigorously evaluate the map \( F_{mN} \) on the computer. The idea is to approximate (4.47) by using a quadrature rule based at the Chebyshev points and to compute a bound for the associated error.

Clenshaw-Curtis quadrature Let \( \left( t_{j}^{M_0} \right)_{j=0}^{M_0} \) denote the Chebyshev points of order \( M_0 \in \mathbb{N}_{\geq 2} \) (see Definition 4.2.6). The parameter \( M_0 \) is the order of the quadrature rule and need not be equal to \( m \). If \( f : [-1, 1] \to \mathbb{R} \) is sufficiently smooth and \( M_0 \) is sufficiently large, then
\[
\int_{-1}^{1} f(s) \, ds \approx \int_{-1}^{1} f_{M_0}(s) \, ds,
\]
where \( f_{M_0} \) is the Chebyshev interpolant of \( f \) of order \( M_0 \). In particular, note that
\[
f_{M_0} = \sum_{j=0}^{M_0} f \left( t_{j}^{M_0} \right) \phi_j,
\]
\[179\]
where \((\phi_j)_{j=0}^{M_0}\) are the Lagrange polynomials associated to \((t_j^{M_0})_{j=0}^{M_0}\), i.e.,

\[
\phi_j(t_l^{M_0}) = \delta_{jl}, \quad 0 \leq j, l \leq M_0.
\]

Hence

\[
\int_{-1}^{1} f_{M_0}(s) \, ds = \sum_{j=0}^{M_0} f(\tau_j) w_j, \quad w_j := \int_{-1}^{1} \phi_j(s) \, ds, \quad 0 \leq j \leq M_0. \tag{4.48}
\]

**Remark 4.7.1.** The weights \((w_j)_{j=0}^{M_0}\) are independent of the objective function \(f\) and are commonly referred to as the Chebyshev quadrature weights. These weights can be efficiently computed by using the Discrete Fourier Transform as explained in [77].

In conclusion, the integral of \(f\) can be approximated by

\[
\sum_{j=0}^{M_0} f(\tau_j) w_j.
\]

This particular quadrature rule is referred to as Clenshaw-Curtis quadrature. Furthermore, the latter approximation is exact whenever \(f\) is a polynomial of at most order \(M_0\). The reader is referred to [68, 77] for a more detailed treatment of Clenshaw-Curtis quadrature.

The next theorem, presented in [68], provides a bound for the error associated to Clenshaw-Curtis quadrature:

**Theorem 4.7.2.** Suppose \(f : [-1, 1] \to \mathbb{R}\) can be analytically extended to the open ellipse \(E_\rho\), for some \(\rho > 1\). If \(f\) is bounded on \(E_\rho\), then

\[
\left| \int_{-1}^{1} f(s) \, ds - \sum_{j=0}^{M_0} w_j f(\tau_j) \right| \leq \frac{64}{15} \rho^{-M_0} \rho^{2} \frac{1}{\rho^2 - 1} \sup_{z \in E_\rho} |f(z)|.
\]

**Proof.** See [68, Theorem 19.3]. \(\square\)

As mentioned before, the strategy is to use Clenshaw-Curtis quadrature to approximate (4.47) and to bound the associated error with the aid of the Theorem 4.7.2. To use this theorem, however, we need one final estimate to bound the integrand in (4.47) on \(E_\rho\):
Lemma 4.7.3. Let $\mu > 1$, then
\[
\sup_{z \in \mathcal{E}_\mu} \left| \frac{t_j + 1}{2} \exp \left( \frac{\lambda_k}{2} (t_j + 1) (1 - z) \right) \psi \left( \frac{t_j + 1}{2} (z + 1) - 1 \right) \right|
\]
\[
\leq \frac{t_j + 1}{2} \exp \left( \frac{\lambda_k}{2} (t_j + 1) \left( 1 + \text{sign} (\lambda_k) (\mu + \mu^{-1}) \right) \right)
\]
\[
\cdot \left( |\psi_0| + \sum_{k=1}^{M} |\psi_k| (\mu^k + \mu^{-k}) \right)
\]
for any $0 \leq k \leq N - 1$ and $0 \leq j \leq m - 1$.

Proof. Let $k \in \mathbb{N}_0$ and recall that $T_k$ is defined by
\[
T_k (\eta (\zeta)) = \eta \left( \zeta^k \right), \quad \zeta \in \mathbb{S}^1,
\]
where $\eta : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ is given by $\eta (\zeta) := \frac{1}{2} (\zeta + \zeta^{-1})$, see Definition 4.2.13. In fact, since $T_k$ is entire, (4.49) must hold for all $\zeta \in \mathbb{C} \setminus \{0\}$. Now, a straightforward computation shows that $\eta (\mathbb{S}_\mu) = \partial \mathcal{E}_\mu$. Hence, by setting $\zeta = \mu e^{i\theta}$ for $\theta \in [0, 2\pi]$, it follows that
\[
\sup_{z \in \partial \mathcal{E}_\mu} |T_k(z)| = \sup_{\zeta \in \mathbb{S}_\mu} |T_k (\eta (\zeta))| \leq \frac{1}{2} (\mu^k + \mu^{-k}) .
\]
Therefore, by the Maximum Modulus Principle,
\[
\sup_{z \in \mathcal{E}_\mu} |T_k(z)| \leq \frac{1}{2} (\mu^k + \mu^{-k}) .
\]
Consequently, since $\left\{ \frac{t_j + 1}{2} (z + 1) - 1 : z \in \mathcal{E}_\mu \right\} \subset \mathcal{E}_\mu$ for any $0 \leq j \leq m - 1$, we conclude that
\[
\sup_{z \in \mathcal{E}_\mu} \left| \psi \left( \frac{t_j + 1}{2} (z + 1) - 1 \right) \right| \leq |\psi_0| + \sum_{k=1}^{M} |\psi_k| (\mu^k + \mu^{-k}) .
\]
Finally, observe that
\[
\sup_{z \in \mathcal{E}_\mu} \left| \exp \left( \frac{\lambda_k}{2} (t_j + 1) (1 - z) \right) \right| = \exp \left( \frac{\lambda_k}{2} (t_j + 1) \left( 1 + \text{sign} (\lambda_k) (\mu + \mu^{-1}) \right) \right)
\]
since the semi-major axis of $\mathcal{E}_\mu$ has length $\mu + \mu^{-1}$, which finishes the proof. \(\Box\)
4.8 Long time integration

The Newton-like map $T$ is only expected to be a contraction whenever the integration time $L$ is sufficiently small. To understand why, consider the bounds $\tilde{Z}_{mN}^1$ and $Z_{\infty N}^1$ constructed in Lemmas 4.6.2 and 4.6.4, respectively. Roughly speaking, these bounds measure the contraction rate of the variation of constants map associated to the finite dimensional system of ODEs obtained after truncating (4.20). Observe that both bounds increase whenever the integration time is increased. Therefore, if we increase $L$, we need to modify other parameters (if possible) to ensure that $\tilde{Z}_{mN}^1$ and $Z_{\infty N}^1$ remain sufficiently small.

It is shown in Lemma 4.6.4 that $Z_{\infty N}^1$ is of the form

$$Z_{\infty N}^1 = \frac{L}{m+1} (1 + \epsilon_1 + \epsilon_2) (\log (m+1) + \pi) C_2(L),$$

where $C_2(L) > 0$ is an increasing function of $L$, $\epsilon_1 > 0$ is a weight to control the interpolation errors of $C^0$-functions and $\epsilon_2 > 0$ is a weight to control the truncation error in phase space (see (4.28)). Hence the only way to ensure that $Z_{\infty N}^1$ remains sufficiently small when $L$ is increased is by increasing the number of interpolation points $m$. Note, however, that an increase in $m$ results in a larger space $X_m^{\begin{array}{c}0 \\ \nu \end{array}}$ which in turn causes the bound $\tilde{Z}_{mN}^1$ to grow. More specifically, in numerical experiments we observed that $\tilde{Z}_{mN}^1$ grows linearly in $m$, see Section 4.9.1. Furthermore, $\tilde{Z}_{mN}^1$ depends (approximately) linearly on $\epsilon_1$, see Lemma 4.6.2 and the construction of the enclosures $I_{kj}$ in Section 4.6.1. Moreover, $\epsilon_1$ is the only parameter which can be varied to control $\tilde{Z}_{mN}^1$ (for fixed $L$ and $m$). Therefore, if we increase $m$, we need to decrease $\epsilon_1$ in order to keep $\tilde{Z}_{mN}^1$ sufficiently small.

In conclusion, whenever we increase the integration time, we need to simultaneously increase $m$ and decrease $\epsilon_1$ in order to ensure that $Z_{\infty N}^1$ and $\tilde{Z}_{mN}^1$ remain sufficiently small. More precisely, in light of the conditions in Theorem 4.3.13, we require that $\tilde{Z}_{mN}^1 < 1$ and $Z_{\infty N}^1 < \epsilon_1$. Furthermore, numerical evidence shows that $\tilde{Z}_{mN}^1$ is approximately of the form $\tilde{Z}_{mN}^1 \approx C_1(L) m \epsilon_1$, where $C_1(L) > 0$ is an increasing function of $L$, see Section 4.9.1. Altogether, the requirements $\tilde{Z}_{mN}^1 < 1$ and $Z_{\infty N}^1 < \epsilon_1$ yield the following constraint:

$$LC_1(L) C_2(L) \frac{m (\log (m+1) + \pi)}{m+1} < \frac{1}{1 + \epsilon_1 + \epsilon_2}.$$  

This constraint explains why $T$ is only expected to be a contraction for sufficiently small $L$. In addition, it shows that $m$ should not be chosen too large. These limitations are explored in more detail for some specific examples in Section 4.9.1.

To perform long time integration we present two options: domain decomposition (see Section 4.8.1) and time-stepping (see Section 4.8.2). As mentioned before, in this paper we primarily focus on time-stepping. Nevertheless, since domain decomposition is a powerful generalization of the single step integration technique and is well-suited for solving boundary value problems, we explain the necessary details in Section 4.8.1. The technicalities are benign and there is an interesting twist in the choice of the approximate inverse $A$ (see Definition 4.8.5).
4.8.1 Domain decomposition

The idea of domain decomposition is to subdivide $[-1, 1]$ into smaller subdomains and to split (4.20) up into a system of infinite dimensional ODEs (one on each subdomain) which are then solved in parallel. The set up for domain decomposition is a straightforward generalization of the theory developed in the previous sections and is similar to the approach in Chapter 2 for finite dimensional systems of ODEs. For these reasons, we shall be brief in the exposition and only present the essential differences.

Let $\{\tau_0 = -1 < \tau_1 < \ldots < \tau_n = 1\}$ be a partition of $[-1, 1]$, where $n \in \mathbb{N}$. Set $a^0 := p$, then (4.20) is equivalent to

$$\left\{ \begin{array}{l}
\frac{d a^i_k}{dt} = \frac{\tau_i - \tau_{i-1}}{2} \left( \lambda_k a^i_k(t) + \frac{L}{2} c_k(a^i(t)) \right), \quad t \in [-1, 1], \\
\lambda_k := \frac{\tau_i - \tau_{i-1}}{2} \lambda_k, \quad 1 \leq i \leq n, \quad k \in \mathbb{N}_0.
\end{array} \right. \quad (4.50)$$

where $1 \leq i \leq n$ and $k \in \mathbb{N}_0$. For notational convenience, set

Then integration of $(P_i)_{i=1}^n$ with the aid of the variation of constants formula yields the following map:

**Definition 4.8.1.** Let $\nu > 1$ and set $X_{\nu,n} := \bigoplus_{i=1}^n X_{\nu}$. For $1 \leq i \leq n$ define $f_i : X_{\nu} \to X_{\nu}$ by

$$[f_i(a^i)]_k(t) := e^{\lambda_k(t+1) a^i_k(1)} + \frac{L(\tau_i - \tau_{i-1})}{4} \int_{-1}^t e^{\lambda_k(s-t)} c_k(a^i(s)) \, ds - a^i_k(t),$$

where $t \in [-1, 1]$ and $k \in \mathbb{N}_0$. The zero finding map $F_n : X_{\nu,n} \to X_{\nu,n}$ for $(P_i)_{i=1}^n$ is defined by $F_n := (f_1, \ldots, f_n)$.

**Proposition 4.8.2.** If $F_n(a) = 0$, where $a = (a^1, \ldots, a^n) \in X_{\nu,n}$, then

$$u(t,x) := \sum_{i=1}^n \mathbf{1}_{\left( \frac{L}{2} (\tau_i+1), \frac{L}{2} (\tau_i+1) \right]}(t) \left( a^i_0(s_i(t)) + 2 \sum_{k=1}^{\infty} a^i_k(s_i(t)) \cos(kx) \right),$$

solves (4.1), where

$$s_i(t) := \frac{2}{\tau_i - \tau_{i-1}} \left( \frac{2t}{L} - \frac{\tau_i - \tau_{i-1} - 1}{L} \right) - 1, \quad 1 \leq i \leq n, \quad t \in [0, L].$$

**Proof.** This is a straightforward generalization of Proposition 4.3.3. \qed

The space $X_{\nu,n}$ is equipped with the max-norm, i.e.,

$$\|a\|_{X_{\nu,n}} := \max_{1 \leq i \leq n} \|a^i\|_{X_{\nu}}, \quad a = (a^1, \ldots, a^n) \in X_{\nu,n},$$

183
where $∥·∥_{X_ν}$ is defined in (4.28). In principle, the weights $\epsilon_1$ and $\epsilon_2$ are allowed to be different on each subdomain. More precisely, for each $1 \leq i \leq n$, we can choose positive weights $(\epsilon_{1,i}, \epsilon_{2,i})$ and then use these in (4.28) to define $∥·∥_{X_ν}$. Strictly speaking then, we should incorporate the dependence of the norm on $(\epsilon_{1,i}, \epsilon_{2,i})$ into the notation. However, to avoid clutter in the notation, we have chosen not to do this.

One can now define a finite dimensional reduction of $F_n$ and a Newton-like operator $T_n$ in almost exactly the same way as in Section 4.3.3. For the sake of completeness, we present the essential details. To keep the coupling between the successive subdomains tractable we work with a single choice of $N$ for all subdomains. Let $m \in \mathbb{N}^n$, $a = (a_1, \ldots, a^n) \in X_{\nu,n}$ and define the Galerkin-projection $\Pi_{m} : X_{\nu,n} \rightarrow X_{\nu,n}$ by

$$\Pi_{m} := \bigoplus_{i=1}^{n} \Pi_{m,i},$$

where $\Pi_{m,i} : X_{\nu,n} \rightarrow X_\nu$ is the Chebyshev projection associated to $a^i$, i.e.,

$$\Pi_{m,i}(a) := \left[ P_{m_i}(a^i_0) \atop \vdots \atop P_{m_i}(a^i_{N-1}) \right],$$

see Definition 4.3.6. Similarly, define $\Pi_i(a) := a^i$, $\Pi_{\infty,N} := \Pi_N - \Pi_{m,N}$, where

$$\left( \Pi_N (a) \right)_k := \begin{cases} a^i_k, & 0 \leq k \leq N - 1, \\ 0, & k \geq N, \end{cases}$$

and set $\Pi^i_\infty := I_{X_\nu} - \Pi^i_N$.

**Definition 4.8.3.** Set $X_{\nu,m} := \Pi_{m,N}(X_{\nu,n})$. The finite dimensional reduction $F_{m,N} : X_{\nu,m} \rightarrow X_{\nu,m}$ of $F_n$ is defined by

$$F_{m,N} := \Pi_{m,N} \circ F_n \circ \Pi_{m,N}|_{X_{\nu,m}}.$$
for $1 \leq i \leq n$, $t \in [-1,1]$ and $k \in \mathbb{N}_0$. We can now use the same reasoning as in Section 4.3.3 to construct an approximation of $DF_n(\hat{a})$ and its inverse. However, there is one crucial difference in the case of domain decomposition. Namely, for $2 \leq i \leq n$, we also have to consider the terms

$$t \mapsto e^{\lambda_{k,i}(t+1)}h_k^{i-1}(1).$$

If $m_i$ is sufficiently large, then the interpolation error associated to (4.51) will be small in comparison to the interpolation error associated to $h_k^i$, since (4.51) is smooth (entire even) and $h_k^i$ is only continuous. Hence we still expect that

$$\Pi_{\infty N}^i DF_n(\hat{a}) \approx -\Pi_{\infty N}^i, \quad 1 \leq i \leq n.$$ 

It is not true anymore, however, that $\|t \mapsto e^{\lambda_{k,i}(t+1)}h_k^{i-1}(1)\|_{\mathcal{X}_\nu^\infty}$ is small in comparison to $\Pi_{\infty N}^i(h)\|_{\mathcal{X}_\nu^\infty}$. Indeed, due to the choice of the uniform norm on $\mathcal{X}_\nu^\infty$ we cannot directly exploit the effect of the exponentials on $h_k^{i-1}(1)$, since we only “observe” the worst-case scenario. More precisely,

$$\|t \mapsto e^{\lambda_{k,i}(t+1)}h_k^{i-1}(1)\|_{\mathcal{X}_\nu^\infty} = \sup_{t \in [-1,1]} \left\|e^{\lambda_{k,i}(t+1)}h_k^{i-1}(1)\right\|_{\nu} = \|h_k^{i-1}(1)\|_{\nu}.$$ 

The above discussion explains why we need to incorporate (4.51) into the definition of the approximate derivative and inverse. To this end, let $1 \leq i \leq n$ and define $E^+_i : \mathcal{X}_{\nu,n} \to \mathcal{X}_\nu^\infty$ by

$$[E^+_i(h)]_k(t) : = (1 - \delta_{iI})e^{\lambda_{k,i}(t+1)}h_k^{i-1}(1) - h_k^i(t),$$

where $t \in [-1,1]$ and $k \geq N$.

**Definition 4.8.4 (Approximate derivative).** The approximate derivative $\widehat{DF}_n : \mathcal{X}_{\nu,n} \to \mathcal{X}_{\nu,n}$ of $DF_n(\hat{a})$ is defined by

$$\widehat{DF}_n : = \bigoplus_{i=1}^n \left[\Pi_N^i DF_{m,N}(\hat{a}) \oplus (-\Pi_{\infty N}^i) \oplus E^+_i\right].$$

To see how to construct a suitable approximate inverse, let $(y^1, \ldots, y^n) \in \bigoplus_{i=1}^n \mathcal{X}_\nu^\infty$ and note that the system of equations $E^+_i(h) = y^i$ for $1 \leq i \leq n$ can be explicitly solved. More specifically, a straightforward computation shows that $h^1_k = -y_k^1$ and

$$h^i_k(t) = -e^{\lambda_{k,i}(t+1)} \sum_{j=1}^{i-1} \exp \left(2 \sum_{l=j+1}^{i-1} \lambda_{k,l}\right) y_k^l(1) - y_k^i(t), \quad 2 \leq i \leq n,$$
where \( k \geq N \) and \( t \in [-1, 1] \). In particular, if \( N \) is sufficiently large, then
\[
h_k^i(t) \approx -e^{\lambda_k i(t+1)} y_k^{i-1}(1) - y_k^i(t), \quad k \geq N, \ t \in [-1, 1].
\]
Motivated by this observation, we define the mappings \( E_i^- : \mathcal{X}_{\nu,n} \to \mathcal{X}_{\nu}^{\infty} \) by
\[
[E_i^-(h)]_k(t) := -(1 - \delta_{i1}) e^{\lambda_k i(t+1)} h_k^{i-1}(1) - h_k^i(t)
\]
for \( 1 \leq i \leq n \).

**Definition 4.8.5** (Approximate inverse). The approximate inverse \( A_n : \mathcal{X}_{\nu,n} \to \mathcal{X}_{\nu,n} \) of \( DF_n (\hat{a}) \) is defined by
\[
A_n := \bigoplus_{i=1}^n \left[ \Pi_N A_{mN} \oplus (-\Pi_{\infty N}) \oplus E_i^- \right].
\]

In the next lemma we show that with the above modifications it is still justified to call \( A_n \) and \( DF_n \) approximate inverses of each other.

**Lemma 4.8.6.** Let \( 1 \leq i \leq n \), then
\[
\| \Pi_{\infty} \left( I - A_n \tilde{DF}_n \right) h \|_{\mathcal{X}_\nu^{\infty}} \leq \begin{cases} 0, & i \in \{1, 2\}, \\ \epsilon_2, & 3 \leq i \leq n, \end{cases}
\]
for any \( h \in \prod_{j=1}^n B_{1,\epsilon_{1,j},\epsilon_{2,j}}(0) \).

**Proof.** It suffices to observe that
\[
\left( \Pi_{\infty} A_n \tilde{DF}_n h \right)_k(t) = -(1 - \delta_{i1}) (1 - \delta_{i-11}) e^{\lambda_k i(t+1)} e^{2\lambda_{k,i-1}} h_k^{i-2}(1)
\]
for any \( t \in [-1, 1] \) and \( k \geq N \).

**Remark 4.8.7.** In order for this bound to be sufficiently small, i.e., strictly below \( \epsilon_{2,i} \), we need \( N \) to be sufficiently large.

Finally, we define a Newton-like operator for \( F_n \):

**Definition 4.8.8.** The Newton-like operator \( T_n : \mathcal{X}_{\nu,n} \to \mathcal{X}_{\nu,n} \) for \( F_n \) based at \( \hat{a} \) is defined by
\[
T_n := I - A_n F_n.
\]

One can prove that \( T_n \) is a contraction by verifying the inequalities in Theorem 4.3.13 component-wise for \( \Pi_{mN}^i T_n, \Pi_{\infty}^i T_n, \Pi_{mN}^i T_n \) and \( 1 \leq i \leq n \). The estimates developed in Sections 4.5 and 4.6 are easily adapted to the domain decomposition setting. The most significant differences occur in the tail-estimates which stem from the additional “boundary terms” in the tail part of the approximate inverse. More precisely, for \( 2 \leq i \leq n \) we need to modify the bounds \( Y_{\infty}, Z_{\infty}^1 \) and \( Z_{\infty}^2 \) in Proposition 4.5.5 and Lemmas 4.6.6 and 4.6.11, respectively. We have outlined the required modifications for \( 2 \leq i \leq n \) below. For notational convenience, we will write
\[
L_i := \frac{L (\tau_i - \tau_{i-1})}{4}, \quad \bar{\epsilon}_i := 1 + \epsilon_{1,i} + \epsilon_{2,i}, \quad 1 \leq i \leq n.
\]
Modifications for $Y_\infty$ A straightforward computation shows that

$$
[\Pi^i_\infty A_n F_n (\hat{a})]_k (t) = -e^{\lambda_k,i (t+1)} \left[ L_{i-1} \int_{-1}^{1} e^{\lambda_{k,(i-1)} (1-s)} c_k (\hat{a}^{i-1}(s)) \, ds + \delta_{i2} e^{2 \lambda_{k,(i-1)}} p_k \right] - L_i \int_{-1}^{t} e^{\lambda_k,i (t-s)} c_k (\hat{a}^i(s)) \, ds
$$

for $2 \leq i \leq n$, $k \geq N$ and $t \in [-1, 1]$. Hence

$$
\| \Pi^i_\infty A_n F_n (\hat{x}) \|_{X^\infty_\nu} \leq \frac{L_{i-1} \left( e^{2 \lambda_{N,(i-1)}} - 1 \right)}{\lambda_{N,(i-1)}} \| \Pi_\infty c (\hat{a}^{i-1}) \|_{X^\infty_\nu} + \delta_{i2} \delta_p + \frac{L_i \left( e^{2 \lambda_{N,i}} - 1 \right)}{\lambda_{N,i}} \| \Pi_\infty c (\hat{a}^i) \|_{X^\infty_\nu}
$$

by Lemma 4.4.2 where $\delta_p$ is an a-priori bound for $\| \Pi_\infty (p) \|_{X^\infty_\nu}$.

Modifications for $Z^1_\infty$ Observe that

$$
[\Pi^i_{\infty} A_n (DF_n (\hat{a}) - \hat{DF}_n) h]_k (t) = -e^{\lambda_k,i (t+1)} L_{i-1} \int_{-1}^{1} e^{\lambda_{k,(i-1)} (1-s)} Dc_k (\hat{a}^{i-1}(s)) h^{i-1}(s) \, ds - L_i \int_{-1}^{t} e^{\lambda_k,i (t-s)} Dc_k (\hat{a}^i(s)) h^i(s) \, ds
$$

for $2 \leq i \leq n$, $k \geq N$ and $t \in [-1, 1]$. Therefore, if $h \in \prod_{j=1}^{n} B_{1,(\epsilon_{i1},\epsilon_{i2})}(0)$, then

$$
\| \Pi^i_{\infty} A_n (DF_n (\hat{a}) - \hat{DF}_n) h \|_{X^\infty_\nu} \leq \frac{L_{i-1} \left( e^{2 \lambda_{N,(i-1)}} - 1 \right)}{\lambda_{N,(i-1)}} \hat{\epsilon}_{i-1} \max_{s \in [-1,1]} \| \Gamma (\hat{g}^{i-1}_1(s)) \|_{B(\ell^1,\ell^1)} + \frac{L_i \left( e^{2 \lambda_{N,i}} - 1 \right)}{\lambda_{N,i}} \hat{\epsilon}_i \max_{s \in [-1,1]} \| \Gamma (\hat{g}^i_1(s)) \|_{B(\ell^1,\ell^1)}
$$

by Lemma 4.4.2 where the coefficients $\{\hat{g}^i_1 : 1 \leq i \leq n\}$ are defined in (4.31).
Modifications for $Z_{\infty}^2$ Define $\eta^i_k : [-1, 1] \to \mathbb{R}$ by

$$\eta^i_k(s) := \int_0^1 D^2 c_k (\hat{a}^i(s) + \tau rv^i(s)) [h^i(s), v^i(s)] \, d\tau, \quad 1 \leq i \leq n,$$

then

$$\left[ \prod_{i=1}^\infty A_n (DF_n (\hat{a} + rv) - DF_n (\hat{a})) h \right]_k (t)$$

$$= - \left( e^{\lambda_{k,i}(t+1)} L_{i-1} \int_{-1}^1 e^{\lambda_{k;i-1}(1-s)} \eta^i_{k-1}(s) \, ds + L_i \int_{-1}^t e^{\lambda_{k;i}(t-s)} \eta^i_k(s) \, ds \right) r$$

for $2 \leq i \leq n, k \geq N$ and $t \in [-1, 1]$ by the computation in (4.46). Hence (as before), if $h \in \prod_{j=1}^n B_{1/(\epsilon_{1,j}, \epsilon_{2,j})}(0)$, then

$$\| \Pi_{\infty}^i A_n (DF_n (\hat{a} + rv) - DF_n (\hat{a})) h \|_{\chi^\infty}$$

$$\leq \frac{L_{i-1} \left( e^{2\lambda_{N,i}(1-i)} - 1 \right)}{\lambda_{N,i}} \tilde{\epsilon}_{i-1} \max_{s \in [-1, 1]} \tilde{g}'' \left( \| \tilde{a}^{i-1}(s) \|_\nu + r^* \tilde{\epsilon}_i \right) r$$

$$+ \frac{L_i \left( e^{2\lambda_{N,i}} - 1 \right)}{\lambda_{N,i}} \tilde{\epsilon}_i \max_{s \in [-1, 1]} \tilde{g}'' \left( \| \tilde{a}^i(s) \|_\nu + r^* \tilde{\epsilon}_i \right) r$$

by Lemmas 4.4.2 and 4.6.8

4.8.2 Time stepping

The idea of time-stepping is to start with some initial condition and to rigorously integrate the associated initial value problem on a small time interval. If the proof is successful, we compute a rigorous enclosure for the endpoint of the orbit. This enclosure is then used as a new initial condition which we try to integrate again. This process is repeated as long as possible.

The latter scheme is only successful if we can compute sharp enclosures for the endpoint of the orbit. To this end, suppose we have successfully proved the existence of a fixed point $a^* \in \hat{a} + B_{\tilde{r},\epsilon}(0)$ of $T$ for some $\tilde{r} > 0$, then

$$a^*_k(1) = e^{2\lambda_k p_k} + \frac{L}{2} \int_{-1}^1 e^{\lambda_k(1-s)} c_k (a^*(s)) \, ds$$

$$= e^{2\lambda_k p_k} + \frac{L}{2} \int_{-1}^1 e^{\lambda_k(1-s)} c_k (\hat{a}(s)) \, ds$$

$$+ \frac{L_{\tilde{r}}}{2} \int_{-1}^1 e^{\lambda_k(1-s)} Dc_k (\hat{a}(s)) \, h(s) \, ds$$

(4.52)

(4.53)
\[
\frac{L\hat{\tau}^2}{2} \int_{-1}^{1} e^{\lambda_k(1-s)} \int_{0}^{1} (1 - \tau) D^2 c_k (\hat{a}(s) + \tau \hat{r}h(s)) [h(s), h(s)] d\tau \, ds,
\]

for any \( k \in \mathbb{N}_0 \) and some \( h \in B_{1,\epsilon}(0) \). We shall use the latter expansion to compute an enclosure for \( a^*(1) \). We stress that this expansion is necessary for obtaining a sufficiently sharp enclosure. In particular, the “trivial” enclosure for \( a^*(1) \) based on the estimate \( \| a^* - \hat{a} \|_{X_{\nu}} \leq \hat{r} \epsilon \), where \( \epsilon = 1 + \epsilon_1 + \epsilon_2 \) as before, is too crude for long time integration (see the discussion about the propagation of errors in Section 4.9.2).

Now, let us start by computing an enclosure for \( (a^*_k(1))_{k=0}^{N-1} \). We remark that the needed computations are very similar to the ones in Lemmas 4.6.2 and 4.6.9.

**Lemma 4.8.9.** There exist intervals \( I_k, J_k \in \mathbb{IF} \), for \( 0 \leq k \leq N-1 \), such that

\[
a^*_k(1) \in e^{2\lambda_k} p_k + \frac{L}{2} \int_{-1}^{1} e^{\lambda_k(1-s)} c_k (\hat{a}(s)) \, ds + I_k + J_k,
\]

for any \( h \in B_{1,\epsilon}(0) \). The formulas for \( I_k \) and \( J_k \) are given in (4.56) and (4.57), respectively.

**Proof.** We start by constructing an interval enclosure for (4.53). Let \( h \in B_{1,\epsilon}(0) \) be arbitrary, then

\[
h(s) \in [- (1 + \epsilon_1), 1 + \epsilon_1] \times \prod_{k=1}^{N-1} \left[ - \frac{1 + \epsilon_1}{2\nu^k}, \frac{1 + \epsilon_1}{2\nu^k} \right] \times \prod_{k=N}^{\infty} \left[ - \frac{\epsilon_2}{2\nu^k}, \frac{\epsilon_2}{2\nu^k} \right],
\]

for any \( s \in [-1,1] \). Recall the definition of \( \mathbf{B}_0 \) in (4.43) and the definitions of \( e_{kj}, L \in \mathbb{IF} \) which were used to construct the intervals \( I_{kj} \), see (4.44). Let \( \mathbf{r} \in \mathbb{IF}, \ h \in \mathbb{IF}^{N_{0}(N-1)} \) be enclosures for \( \hat{r} \) and

\[
[- (1 + \epsilon_1), 1 + \epsilon_1] \times \prod_{k=1}^{N-1} \left[ - \frac{1 + \epsilon_1}{2\nu^k}, \frac{1 + \epsilon_1}{2\nu^k} \right] \times \prod_{k=N}^{\infty} \left[ - \frac{\epsilon_2}{2\nu^k}, \frac{\epsilon_2}{2\nu^k} \right],
\]

respectively (note that the definition of \( h \) is here slightly different than the one defined in Section 4.6.1). Then the same arguments which were used to construct the enclosures \( I_{kj} \) show that in order to bound (4.53) it suffices to set \( I_k := [- I_k^+, I_k^+] \), where

\[
I_k^+ := \triangle \left( L \hat{\tau} \triangle \left( e_{k0} \hat{B}_0 h \right) \right).
\]

Finally, we compute an enclosure for (4.54). To accomplish this, we use a more crude estimate. Namely, observe that

\[
\left| \frac{L\hat{\tau}^2}{2} \int_{-1}^{1} e^{\lambda_k(1-s)} \int_{0}^{1} (1 - \tau) D^2 c_k (\hat{a}(s) + \tau \hat{r}h(s)) [h(s), h(s)] d\tau \, ds \right|
\leq \frac{L\hat{\tau}^2}{4} \epsilon^2 \epsilon_k \max_{s \in [-1,1]} \tilde{g}''(\|\hat{a}(s)\|_{\nu} + \hat{r} \epsilon) \int_{-1}^{1} e^{\lambda_k(1-s)} ds
\]

189
by Lemma 4.6.8 where $\tilde{\epsilon} = 1 + \epsilon_1 + \epsilon_2$ as before. Furthermore, we can use interval arithmetic to compute floating point numbers

$$J_k^+ := \triangle \left( \frac{L \tilde{r}^2}{4} \tilde{\epsilon}^2 \varepsilon_k \max_{s \in [-1,1]} \tilde{g}''(\|\hat{a}(s)\|_\nu + \tilde{r}\tilde{\epsilon}) \right)$$

for $0 \leq k \leq N - 1$. Therefore, to bound (4.54), it suffices to set $J_k := [-J_k^+, J_k^+]$.

Finally, we compute a bound for the tail of $a^*(1)$:

**Lemma 4.8.10.** Suppose that $\|C(\hat{p})\|_\nu \leq \delta_p$, then

$$\left\| (a_k^*(1))_{k=N}^\infty \right\|_\nu \leq e^{2\lambda N} \delta_p + \frac{L (e^{2\lambda N} - 1)}{2\lambda N} \max_{s \in [-1,1]} \tilde{g}''(\|\hat{a}(s)\|_\nu + \tilde{r}\tilde{\epsilon})$$

for any $h \in B_{1,\epsilon}(0)$.

**Proof.** The same arguments as in Proposition 4.5.5 (but now with $t = 1$) show that

$$\left\| \left[ e^{2\lambda_k p_k} + \frac{L}{2} \int_{-1}^1 e^{\lambda_k (1-s)} c_k (\hat{a}(s)) \, ds \right]_{k=N}^\infty \right\|_\nu$$

$$\leq e^{2\lambda N} \delta_p + \frac{L (e^{2\lambda N} - 1)}{2\lambda N} \max_{s \in [-1,1]} \tilde{g}''(\|\hat{a}(s)\|_\nu + \tilde{r}\tilde{\epsilon})$$

Furthermore, we observe that (4.53) reduces to

$$\left[ \frac{L}{2} \int_{-1}^1 e^{\lambda_k (1-s)} Dc_k (\hat{a}(s)) h(s) \, ds \right]_{k=N}^\infty = \left( \Pi_\infty A \left( DF (\hat{a}) - \overline{DF} \right) h \right)(1),$$

and the $\ell^1_\nu$-norm of this term is bounded by the estimate in Lemma 4.6.6. Finally, a similar computation as in Lemma 4.6.11 shows that

$$\left\| \left[ \frac{L \tilde{r}^2}{2} \int_{-1}^1 e^{\lambda_k (1-s)} \int_0^1 (1 - \tau) D^2 c_k (\hat{a}(s) + \tau \tilde{r} h(s)) [h(s), h(s)] \, d\tau \, ds \right]_{k=N}^\infty \right\|_\nu$$

$$\leq \frac{L (e^{2\lambda N} - 1)}{4\lambda N} \tilde{r}^2 \varepsilon_2 \max_{s \in [-1,1]} \tilde{g}''(\|\hat{a}(s)\|_\nu + \tilde{r}\tilde{\epsilon}) .$$

Altogether, this yields the desired bound for the tail of $a^*(1)$. \qed

190
4.9 Applications

In this section we test the performance of the method. Since it involves a relatively large number of “free” parameters, we will start by performing various experiments on a toy-example (Fisher’s equation) in Section 4.9.1. The purpose of these experiments is to get some insight into the dependence of the method on the integration time and the number of interpolation nodes. Next, in Section 4.9.2 we consider a fourth order equation (the Swift-Hohenberg equation) and examine the performance of the method for long time integration.

In order to structure the experiments we have fixed many parameters from the start and only varied the ones we deemed to be relevant. For the sake of clarity, let us recall which parameters we need to choose:

- \( N \in \mathbb{N}_{\geq 2} \): the number of Fourier coefficients in the Galerkin projection,
- \( m \in \mathbb{N} \): the degree of the Chebyshev interpolants for \( a_0, \ldots, a_{N-1} \),
- \( L > 0 \): the integration time (which may be varied at each integration step),
- \( \beta_0 \in \mathbb{R} \): a free parameter in the variation of constants formulation, see Section 4.3.1,
- \( \epsilon_1 > 0 \): a weight for controlling the interpolation errors of \( C^0 \)-functions,
- \( \epsilon_2 > 0 \): a weight for controlling the truncation errors in phase space,
- \( d \in \mathbb{N} \): the number of derivatives used in the computation of a bound for the interpolation error in Proposition 4.5.1,
- \( \rho > 1 \): the size of the Bernstein-ellipse \( E_{\rho} \) used in the computation of a bound for the interpolation error in Proposition 4.5.1,
- \( M_0 \in \mathbb{N} \): the number of Chebyshev points used in the approximation of (4.47) with Clenshaw-Curtis quadrature,
- \( \mu > 1 \): the size of the Bernstein-ellipse \( E_{\mu} \) used in the computation of a bound for the quadrature errors in Lemma 4.7.3,
- \( \nu > 1 \): a lower bound for the decay rate of the Fourier (cosine) coefficients,
- \( r^* > 0 \): an a-priori upper bound for the radius \( r > 0 \) (to be determined) on which \( T \) is a contraction, see Lemma 4.6.8.

In both applications we start by fixing \( N \) and \( \epsilon_2 \). These parameters are used to control the truncation errors in phase space. To be more precise, observe that the residual \( Y_\infty \) is small whenever \( N \) is sufficiently large by Proposition 4.5.5. Furthermore, the bound \( Z_\infty^1 \) in Lemma 4.6.6, which determines the contraction rate of \( T \) on \( X_\nu^\infty \), can be controlled by \( N \) as well. For these reasons, we have based the choice of \( N \) on \( \epsilon_2 \). That is, we first fix \( \epsilon_2 \), and then choose \( N \) sufficiently large so that \( Z_\infty^1 < \epsilon_2 \). In both applications we set \( \epsilon_2 = \frac{1}{5} \). This choice was based on
experimentation and the observation that the results did not significantly improve when $\epsilon_2$ was decreased but did deteriorate when it was increased.

Next, we fix the parameters $d, \rho, \mu$ and $M_0$. These parameters are used to manage the bounds for the residuals on $X^N \oplus X^\infty_N$. More specifically, Theorem 4.7.2 shows that the quadrature errors can be made arbitrarily small by choosing $M_0$ sufficiently large (for fixed $\mu > 1$). Furthermore, Proposition 4.5.1 shows that the interpolation errors can be controlled by choosing $m$ and $d$ sufficiently large (relative to $L$). In both applications we fixed $\mu = 2.5, \rho = 10, M_0 = 64$ and $d = m$. This yielded sufficiently small bounds for the residuals. In particular, the quadrature and interpolation errors were near machine precision in each experiment.

Recall that the parameter $r^*$ is used in the computation of the second order bounds. In both applications we fixed $r^* = 10^{-2}$. This choice was partially based on the observation that the bound in Lemma 4.6.8 did not significantly decrease when $r^*$ was decreased any further. Another reason for choosing this particular value was to integrate as long as possible with the aid of time-stepping. In practice, we typically observed that the time-stepping algorithm was about to fail if the validation radius $r$ reached a value of order $10^{-2}$.

Finally, in both applications we chose $\nu$ close to 1, namely, we set $\nu = 1.0001$. The main motivation for this choice was that when $\nu$ was increased the $Y$-bounds deteriorated while the $Z$-bounds remained effectively unchanged. The reason for not setting $\nu = 1$ is that we used the assumption $\nu > 1$ in Proposition 4.3.3 to prove that zeros of $F$ correspond to solutions of the original PDE which are analytic in space. The dependence of the method on the remaining parameters $m, L, \epsilon_1$ and $\beta_0$ is studied in the following sections. More precisely, in Section 4.9.1 we examine the dependence on $m, L, \epsilon_1$ and in Section 4.9.2 we thoroughly examine the dependence on $\beta_0$. In particular, we have investigated (heuristically) optimal choices for the parameters. We have not, however, made any attempt to construct an algorithm to systematically optimize the parameters.

### 4.9.1 Fisher’s equation

In this section we consider Fisher’s equation, see (4.2). We will validate solutions of (4.2) with the aid of time stepping, as explained in Section 4.8.2 where all parameters are fixed throughout the integration procedure. In particular, we set $N = 10$ in each experiment.

**Initial data** In order to provide some context for the experiments in this section we start by giving a brief summary of the dynamics of Fisher’s equation. To this end, recall that the equation in (4.2) can be transformed into an ODE on $\ell_1^\nu$ via a Fourier cosine transformation in space, see (4.20). We will consider the associated semi-flow on $\ell_1^\nu$, generated by

$$
\frac{d a_k}{d t} = (1 - k^2) a_k - (a * a)_k, \quad k \in \mathbb{N}_0.
$$

(4.58)
It is well known that the latter dynamical system has a sink at $[1 \ 0 \ 0 \ \ldots ]^T \in \ell_1^\nu$ corresponding to $u \equiv 1$.

Throughout this section we consider the initial condition

$$p \approx [1.4054 \ 0 \ -0.3250 \ 0 \ -0.04117 \ 0 \ -0.02214 \ 0 \ -0.01596 \ 0]^T$$

with $\delta_p = 0$, which are (approximately) the first ten Fourier cosine coefficients of the Gaussian function $f(x) = 2 \exp\left(-\frac{1}{2}(x - \frac{\pi}{2})^2\right)$ centered at $x = \frac{\pi}{2}$. The reader is referred to the code available at [67] for the exact initial data.

**Short time integration: one time step**

The goal in this section is to study the largest feasible integration time for one time step as a function of $m$ and $\epsilon_1$. Clearly, this integration time depends heavily on the initial condition and the specific PDE under consideration. Nevertheless, we have tried to construct a representative example which illustrates the typical behavior of the method. In order not to complicate matters, we have set $\beta_0 = 0$ and postponed a thorough analysis of this parameter to the following sections.

We start by examining the dependence of the bound $\tilde{Z}_{mN}$ in Lemma 4.6.2, which determines the contraction rate of $T$.

**Dependence of $\tilde{Z}_{mN}$ on $m$** In this experiment we fixed the integration time to $L = 0.1$. Numerical experiments suggest that the qualitative results in this paragraph are independent of $L$. More precisely, we repeated the experiment for various values of $L \in (0, 0.5]$, which yielded the same qualitative results.

Recall that $\tilde{Z}_{mN}$ depends linearly on $\epsilon_1$ (approximately at least), see the construction of $\tilde{Z}_{mN}$ in Section 4.6.1. To emphasize the dependence on $\epsilon_1$, we will write $\tilde{Z}_{mN} = Z_{mN}(\epsilon_1)$. We have investigated the dependence of $\tilde{Z}_{mN}(1)$ on $m$ by computing the points $\left\{(m, \tilde{Z}_{mN}(1)) : m = 5(j + 1), 1 \leq j \leq 9\right\}$. The result is shown in Figure 4.3 and indicates that the dependence of $\tilde{Z}_{mN}$ on $m$ is linear as well. Altogether, based on the numerical evidence, we conclude that $\tilde{Z}_{mN}$ is (approximately) of the form $\tilde{Z}_{mN} \approx C_1(L)m\epsilon_1$, where $C_1(L) > 0$ is an increasing function of $L$.

**Dependence of $L_{max}$ on $m$ and $\epsilon_1$** Next, we examine the largest feasible integration time as a function of $m$ and $\epsilon_1$. For this purpose, we subdivided the interval $[0.001, 0.3]$ into 300 evenly distributed points $(\epsilon_{1,j})_{j=1}^{300}$. We then fixed $\epsilon_1 = \epsilon_{1,j}$, for some $1 \leq j \leq 300$, and tried to validate the initial value problem for $L = 0.01$ and $m = 10$. If the proof was successful, we increased the integration time by 0.005 and repeated the validation process. If the proof was not successful, we increased $m$ by 5 and tried to validate again. An integration time $L$ was classified as unfeasible if
the validation was unsuccessful for all \( m \in \{5l : 2 \leq l \leq 10 \} \). The maximal integration time \( L_{\text{max}} = L_{\text{max}}(m, \epsilon_1, j) \) was defined as the largest feasible integration time.

The experiment was performed for \( \epsilon_1 \in \{\epsilon_{1,j} : 1 \leq j \leq 300 \} \). The results are shown in Figure 4.3. To explain the results, we first focus on the region where \( \epsilon_1 \) is relatively small (i.e. \( \epsilon_1 \in [0.001, 0.1] \)). The validation failed for \( \epsilon_1 \in [0.001, 0.01] \), since the number of interpolation points needed to ensure that \( Z_{1,\infty,N}^1 < \epsilon_1 \) was too large (i.e. \( m > 50 \)). Next, consider the interval on which \( L_{\text{max}} = 0.005 \), which for notational convenience we shall denote by \( [\epsilon_{1,j_0}, \epsilon_{1,j_1}] \). For \( \epsilon_1 \) close to \( \epsilon_{1,j_0} \), the number of interpolation points needed to ensure that \( Z_{1,\infty,N}^1 < \epsilon_1 \) was relatively large. As \( \epsilon_1 \) increased, this number decreased until the point at which a larger integration time became feasible. When the integration time increased (i.e. jumped from 0.005 to 0.01), the bound \( Z_{1,\infty,N}^1 \) increased as well and in turn caused the needed number of interpolation points to increase again. This process was repeated every time the integration time increased and explains the results for \( \epsilon_1 \in [0.001, 0.1] \).

Finally, consider the region where \( \epsilon_1 \) is relatively large (i.e. \( \epsilon_1 \in [0.1, 0.3] \)). In this region the number of interpolation points needed to ensure that \( Z_{1,\infty,N}^1 < \epsilon_1 \) was relatively small. This allowed for relatively large integration times with a small number of interpolation points. However, when \( \epsilon_1 \) became “too large” (\( \epsilon_1 > 0.23 \)), we needed both \( m \) and \( L_{\text{max}} \) to be sufficiently small in order to keep control over \( \tilde{Z}_{m,N}^1 \). This is the reason why \( L_{\text{max}} \) and \( m \) eventually started to decrease. In conclusion, for small \( \epsilon_1 \) the bottleneck of the method is determined by \( Z_{1,\infty,N}^1 \) and for large \( \epsilon_1 \) by \( \tilde{Z}_{m,N}^1 \).
Figure 4.4: The largest feasible integration time $L_{\text{max}}$ as a function of $m$ and $\epsilon_1$. The sudden “dips” are an artifact of the choice to discretize $m$ in steps of 5.

**Long time integration: multiple time steps**

To illustrate the effectiveness of time-stepping we have integrated (4.2) for $\tau = 8$ time units by using 400 time steps of length $L = 0.02$, $m = 15$ and $\epsilon_1 = 0.125$. These parameter values were based on the results in Figure 4.4. The validated orbit is depicted in Figure 4.1 which indicates that the orbit converges to the trivial equilibrium state $u \equiv 1$. We have carried out the integration for $\beta_0 = 0$ (the “test-case” in which the extra degree of freedom is not used) and $\beta_0 = -1$. The choice $\beta_0 = -1$ was motivated by the observation that the orbit seemed to converge relatively quickly to the homogenous equilibrium state $u \equiv 1$. In particular, near this homogenous equilibrium state, we have that $Dc(\hat{a})h \approx -h$ (for $\beta_0 = 0$). Therefore, when the orbit is near this equilibrium, we can “eliminate” the term $Dc(\hat{a})h$ by setting $\beta_0 = -1$. The expectation is that this choice improves the performance of the method (at least near the equilibrium).

The $C^0$-error between the exact and approximate orbit at each integration step is shown in Figure 4.5. The results show that the choice $\beta_0 = -1$ yields a significant improvement in the performance of the method. Indeed, the $C^0$-error was decreasing at a much faster rate for $\beta_0 = -1$. In fact, for $\beta_0 = 0$ the $C^0$-error became almost constant as the orbit approached the equilibrium, whereas the error for $\beta_0 = -1$ kept decreasing. To get some more insight into these results, we have depicted the evolution of the radii $\hat{r}_{p_0}$ and $\hat{r}_{p_1}$ of the enclosures around the zeroth and first Fourier coefficient, respectively, in Figure 4.6 (see Lemma 4.8.9). We observe that the size of the enclosures for $\beta_0 = -1$ decreased at a much faster rate
Figure 4.5: A rigorous error bound for the $C^0$-error between the exact and approximate solution $u^*$ and $\hat{u}$, respectively, at each integration step on a semi-logarithmic scale for $\beta_0 \in \{-1, 0\}$.

than for $\beta_0 = 0$. In fact, for $\beta_0 = 0$ the radius $\hat{r}_{p_0}$ was increasing at each integration step, which was caused by the fact that $\lambda_0 = 0$ for $\beta_0 = 0$. In turn, this explains why the $C^0$-error eventually stopped improving for $\beta_0 = 0$. The reader is referred to Section 4.9.2 for a more detailed discussion about the relationship between the parameters $(\lambda_k)_{k=0}^{N-1}$ and the rate at which the enclosures grow. Altogether, the results indicate that one could (in principle) integrate forward indefinitely with the choice $\beta_0 = -1$, but not with $\beta_0 = 0$.

### 4.9.2 Swift-Hohenberg

In this section we consider the Swift-Hohenberg equation to illustrate the effectiveness of our method for more complicated PDEs. The objective is to examine the dependence of the method on the parameter $\beta_0$ by performing long time integration (through time-stepping) for qualitatively different choices of $\beta_0$. To accomplish this, we fix all other parameters. In particular, we set $N = 5$, $m = 20$, $\epsilon_1 = 0.15$ and $L = 0.005$. These parameter values were obtained by performing a similar experiment as in Section 4.9.1 and choosing “optimal” parameter values. Note that

$$\lambda_k = \frac{L}{2} \left( \beta_0 + k^2 \left( 2 - k^2 \right) \right), \quad k \in \mathbb{N}_0,$$
for a time step of length $L$. To emphasize the dependence of $\lambda_k$ on $\beta_0$ we shall frequently write $\lambda_k = \lambda_k(\beta_0)$.

**Initial data** The semi-flow on $\ell^1_\nu$ induced by (4.3) is generated by

$$\frac{da_k}{dt} = (k^2 (2 - k^2) + r - 1) a_k + (a * a)_k - (a * a * a)_k, \quad k \in \mathbb{N}_0. \quad (4.60)$$

Observe that the origin is an equilibrium and that there exist two other trivial equilibrium states $u \equiv a_0 = \frac{1}{2} (1 \pm \sqrt{4r - 3})$ if $r > \frac{3}{4}$. If $r \in (0, 1)$, then the origin is a saddle with an one dimensional unstable manifold and $T_0 W^u(0) = \text{span} \left( \begin{bmatrix} 0 & 1 & 0 & \infty \end{bmatrix}^T \right)$.

Throughout this section we set $r = \frac{4}{5}$ and consider a (fixed) initial condition close to a point on $W^u(0)$. The reason for choosing such an initial condition is to make the existence of a connecting orbit from the origin to another equilibrium plausible. For this purpose, we set $p := \begin{bmatrix} 0 & 0.1 & 0 \infty \end{bmatrix}^T$. In numerical experiments we observed that the associated orbit converged to a nontrivial equilibrium state, see Figure 4.2. A future research project would be to make this assertion rigorous and to prove the existence of a connecting orbit by combining the method in this paper with the techniques developed in [63].

**Variation of $\beta_0$** Recall that the parameter $\beta_0$ was introduced to “redistribute” the linear term $u$ amongst the differential operator $\mathcal{L}$ and the nonlinearity $g$, see Section 4.3. In particular, there are two “special” cases: $\beta_0 = r - 1 = -\frac{1}{5}$ and $\beta_0 = 0$. The choice $\beta_0 = r - 1$ corresponds to the case in which the linear term $u$ is completely incorporated into $\mathcal{L}$. The other extreme, $\beta_0 = 0$, corresponds to the
case in which the linear term $u$ is completely incorporated into $g$. The remaining cases, $\beta_0 \in \mathbb{R} \setminus \{r - 1, 0\}$, correspond to a situation in which the linear term $u$ is distributed amongst $\mathcal{L}$ and $g$.

We have varied $\beta_0$ in order to investigate the effect of the sign of $\lambda_0$ and $\lambda_1$ on the performance. To give a more precise explanation, first note that $\lambda_0 = \beta_0$ and $\lambda_1 = \beta_0 + 1$. Therefore, for any choice of $\beta_0 \in \mathbb{R}$, exactly one of the following three cases holds:

- $\lambda_0, \lambda_1 > 0$,
- $\lambda_0 \leq 0, \lambda_1 > 0$,
- $\lambda_0 < 0, \lambda_1 \leq 0$.

In particular, all three cases occur when $\beta_0$ is varied in $(-\infty, 1]$. Furthermore, $\lambda_k < 0$ for $k \in \mathbb{N}_{\geq 2}$ and $\beta_0 \in (-\infty, 1]$. Based on these observations we decided to vary $\beta_0 \in \Delta := \left\{-2 + \frac{j}{5} : 0 \leq j \leq 15\right\}$, since this covers the above three cases (including the “special” ones) and provided a sufficiently detailed picture of the qualitatively different behavior of the method.

Results  
For each $\beta_0 \in \Delta$ we tried to validate the initial value problem in (4.3) for $\tau = \tau_{\text{max}} = 10$ time units by using 2000 time steps of length $L = 0.005$. The choice for $\tau_{\text{max}}$ was motivated by the observation that the corresponding orbit seemed to have converged to a nontrivial equilibrium state, see Figure 4.2. The results are shown in Figure 4.7. The results show that the performance of the method significantly improved when $\beta_0$ was decreased. More precisely, the parameter values $\beta_0 \in [0, 1]$ yielded the “best” results in terms of accuracy and maximal integration time. The optimal choice was $\beta_0 = -1.8$, i.e., for this parameter value the $C^0$-error was minimal and $\tau = \tau_{\text{max}}$. Note that in this case $\lambda_0, \lambda_1 < 0$.

We claim that enforcing $\lambda_0$ and $\lambda_1$ to be negative (by decreasing $\beta_0$) is beneficial and is the reason why the results improved when $\beta_0$ was decreased. To understand why, first observe that a crucial factor that determines the accuracy of the method, and in turn the maximal number of successful integration steps, is the size of the enclosure around the endpoint of the orbit at each integration step. For instance, for $\beta_0 \in [0, 1]$, the sizes of the enclosures around the endpoints were increasing too fast. This caused the bounds for the residuals ($Y_{mN}$ and $Y_{\infty}$) to increase at a fast rate as well and ultimately resulted in a premature breakdown of the validation process.

To explain in more detail how the size of the enclosures grow, let $\hat{r}_{pk}$ denote the radius of the enclosure around the endpoint of the $k$-th Fourier coefficient for $0 \leq k \leq N - 1$ (see Lemma 4.8.9). The sizes of these enclosures are the bottleneck for performing long time integration. Indeed, it follows from Lemma 4.8.10 that the bound for the tail of the endpoint can be controlled whenever $N$ is sufficiently large. However, the only “control” we have over the enclosures for the lower order
Figure 4.7: Results for the validated integration of (4.3) by using a maximum of 2000 time steps (with identical parameters) of length $L = 0.005$: (a) The final integration time $\tau$ as function of $\beta_0 \in \Delta$. (b) A rigorous error bound for the $C^0$-error between the exact and approximate solution $u^*$ and $\hat{u}$, respectively, on $[0, \tau(\beta_0)] \times [0, \pi]$ on a semi-logarithmic scale for $\beta_0 \in \Delta$.

Fourier coefficients are the integration time and the factors $e^{\lambda k}$, see (4.52), (4.53) and (4.54). In particular, these formulae show that enforcing $\lambda_k$ to be negative for all $0 \leq k \leq N - 1$ could potentially aid in “controlling” the rate at which $\hat{r}_{pk}$ grows. Furthermore, note that if $\lambda_k > 0$, then $\hat{r}_{pk}$ will always grow at each integration step due to the term $e^{2\lambda k} p_k$ and wrapping, see (4.52). These observations substantiate the claim that enforcing $\lambda_0, \lambda_1 < 0$ could potentially improve the performance of the method. We have provided numerical evidence for this claim by depicting the evolution of the radii $\hat{r}_{p_0}$ and $\hat{r}_{p_1}$ in Figure 4.8 for the following three cases:

(i) The optimal case $\beta_0 = -1.8$ in which $\lambda_0, \lambda_1 < 0$.

(ii) The case $\beta_0 = -0.6$ in which $\lambda_0 < 0$, $\lambda_1 > 0$. This is the parameter value at which the accuracy of the method starts to deteriorate, see Figure 4.7b.

(iii) The “worst” case $\beta_0 = 1$ in which $\lambda_0, \lambda_1 > 0$.

The results in Figure 4.8 show that (on average) the rates at which $\hat{r}_{p_0}$ and $\hat{r}_{p_1}$ grew decreased as $\beta_0$ decreased. In the first case, $\hat{r}_{p_0}$ and $\hat{r}_{p_1}$ even started to decrease (eventually), which was made possible by having $\lambda_0, \lambda_1 < 0$ and was related to the fact that the orbit converged to a nontrivial equilibrium state. In the second case, the rate at which $\hat{r}_{p_0}$ grew was initially decreasing, but started to increase again at the end. The reason for this was that $\hat{r}_{p_1}$ increased at a relatively large rate throughout the whole integration procedure, which in turn was caused by $\lambda_1$ being positive. In particular, there was a critical integration step where $\hat{r}_{p_1}$ became significantly larger than $\hat{r}_{p_0}$, thereby becoming the bottleneck of the method, which caused $\hat{r}_{p_0}$ to suddenly increase again at a relatively fast rate.
Figure 4.8: The evolution of the radii $\hat{r}_{p_0}$ and $\hat{r}_{p_1}$ of the enclosures around the zeroth and first Fourier-coefficient, respectively, on a semi-logarithmic scale for three qualitatively different choices of $\beta_0$.

Based on the above observations it is tempting to choose $\beta_0 < 0$ extremely negative so that $e^{\lambda_k} \approx 0$ for all $k \in \mathbb{N}_0$. This would certainly improve the bound $\tilde{Z}_{1mN}$ and the size of the enclosures, provided the integration time $L$ is not too small. However, the bounds associated to the interpolation and quadrature errors would deteriorate if $\beta_0$ is decreased too much. Indeed, if $|\beta_0|$ is too large, then (roughly speaking) all the terms in $Y_{\infty N}$ which involve $|\lambda_k|$ would deteriorate, see Proposition 4.5.1. The same holds for the quadrature errors in Lemma 4.7.3. Moreover, the contribution of the linear term in $c(a)$ would be relatively large as well, see (4.19). In turn, this would cause an increase in $\max_{s \in [-1,1]} \| Dc(\hat{a}(s)) \|_{\nu}^*$ and hence deterioration of the bound $Z_{1\infty N}$ (see Lemma 4.6.4).

Finally, we end this section with a suggestion for a future research project. Recall that in Section 4.9.1 we were able to enhance the performance of the method near an equilibrium state by choosing $\beta_0$ in such a way that $Dc(\hat{a})$ was effectively eliminated. This was possible because the equilibrium state was homogenous. In the case of a non-homogenous equilibrium state, such as the one encountered in this section, we cannot completely eliminate $Dc(\hat{a})$. Nevertheless, one can still try to minimize the effect of this term by only “eliminating” it from the finite part. More precisely, instead of using one uniform shift to modify all the eigenvalues of $L$, one could use a finite number of shifts, which may vary per component, to replace a finite number of eigenvalues, see Remark 4.3.2. Then, in the case that $\Pi_N Dc(\hat{a}) \Pi_N$ is diagonal, we would simply shift the eigenvalues in such a
way that all the terms on the diagonal are eliminated. In the general case, we would first need to diagonalize $\Pi_N Dc(\hat{a}) \Pi_N$ by using an appropriate coordinate transformation in the finite part.
Summary

In this dissertation we develop generally applicable validated numerical methods for the computation of invariant objects in differential equations. This includes the validated computation of periodic orbits, (un)stable manifolds and connecting orbits in nonlinear ODEs, as well as the validation of solutions of parabolic semilinear PDEs. The techniques developed in this dissertation are all based on Chebyshev approximations, which constitute a non-periodic analog of Fourier cosine approximations, and parameterized Newton-Kantorovich methods. A brief introduction into the field of validated numerical methods and a general outline of the main ideas used in this dissertation are given in Chapter 1.

In Chapter 2 we present a rigorous numerical method for validating analytic solutions of nonlinear ODEs by using Chebyshev series and domain decomposition. The idea is to define a Newton-like map, whose fixed points correspond to solutions of the ODE, on the space of geometrically decaying Chebyshev coefficients, and to use the so-called radii-polynomial approach to prove that the map has an isolated fixed point in a small neighborhood of a numerical approximation. The novelty of the proposed method is the use of Chebyshev series in combination with domain decomposition. In particular, a heuristic procedure based on the theory of Chebyshev approximations for analytic functions is presented to construct efficient grids for validating solutions of boundary value problems.

In Chapter 3 we present a computer-assisted procedure for proving the existence of transverse heteroclinic orbits connecting hyperbolic equilibria of polynomial vectorfields. The idea is to compute high-order Taylor approximations of local charts on the (un)stable manifolds by using the parameterization method and to use Chebyshev series to parameterize the orbit in between. The existence of a heteroclinic orbit can then be established by setting up an appropriate fixed-point problem amenable to computer-assisted analysis. The novelty of the proposed method is that the (un)stable manifolds and connecting orbit in between are validated simultaneously.

Finally, in Chapter 4 we present a computer-assisted procedure for validating solutions of (scalar) semilinear parabolic PDEs, which can be extended periodically in space. The main idea is to recast the problem into an equivalent zero finding problem on a space of time dependent Fourier coefficients with geometric decay by using the variation of constants formula. We construct a finite dimensional reduction of the zero finding problem by approximating the dynamics of

203
a finite number of Fourier coefficients with the aid of Chebyshev interpolation in
time. Numerical simulation and analysis on paper are then combined to set up an
equivalent fixed point problem by constructing a Newton-like map. Finally, a finite
number of inequalities are derived, which (if satisfied) prove that the Newton-like
map is a contraction in a small neighborhood of the numerical approximation.
Acknowledgments

First of all, I would like to thank my supervisor Jan Bouwe van den Berg for giving me the opportunity to be part of a very exciting field of mathematics. I have enjoyed our collaboration during the last four (and a half) years. From the start, as a Master’s student, I noticed that you had a great deal of trust in my work and gave me the freedom to pursue my own ideas. For that I am very thankful.

I thank the PhD students, postdocs and staff members at the math department of the VU for a nice work environment. I thank Christian Reinhardt for the many math (and non-math) related conversations. I very much enjoyed our discussions about the computation of connecting orbits, which contributed to the work in Chapter 3.

I have had many interesting discussions with Maxime Breden about the validated integration of PDEs for which I would like to thank him.

I would like to thank Jason Mireles-James for his great hospitality. I have had a great time during my research visits in Florida and I enjoyed our collaboration. Although we did not finish solving the (very challenging) problem of validating connecting orbits in parabolic PDEs, it was a very exciting project to be part of and it motivated the work in Chapter 4. Also, the ideas for the delay integrator in Chapter 1 would not have been in this dissertation were it not for our collaboration together with Chris Groothedde.

Last but certainly not least, I would like to thank my parents and sister for their continuous support and encouragement for pursuing a PhD. I am especially thankful to my sister, who does not share my passion for mathematics, but nevertheless endured and listened to me talking about it on many occasions.
Bibliography


